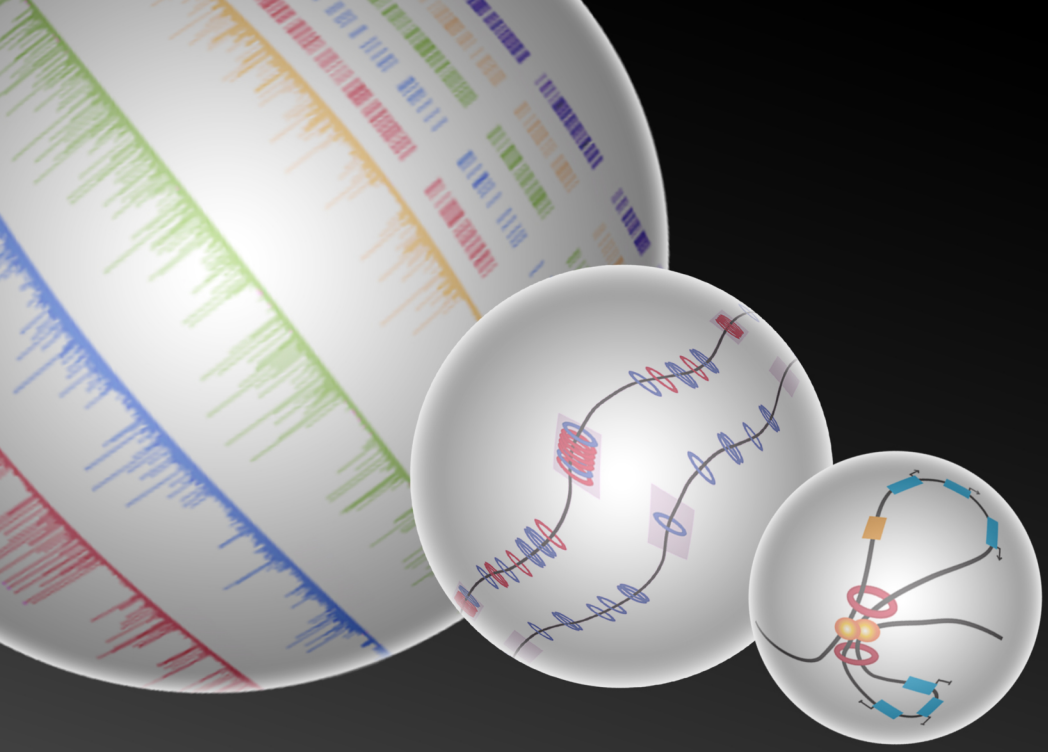
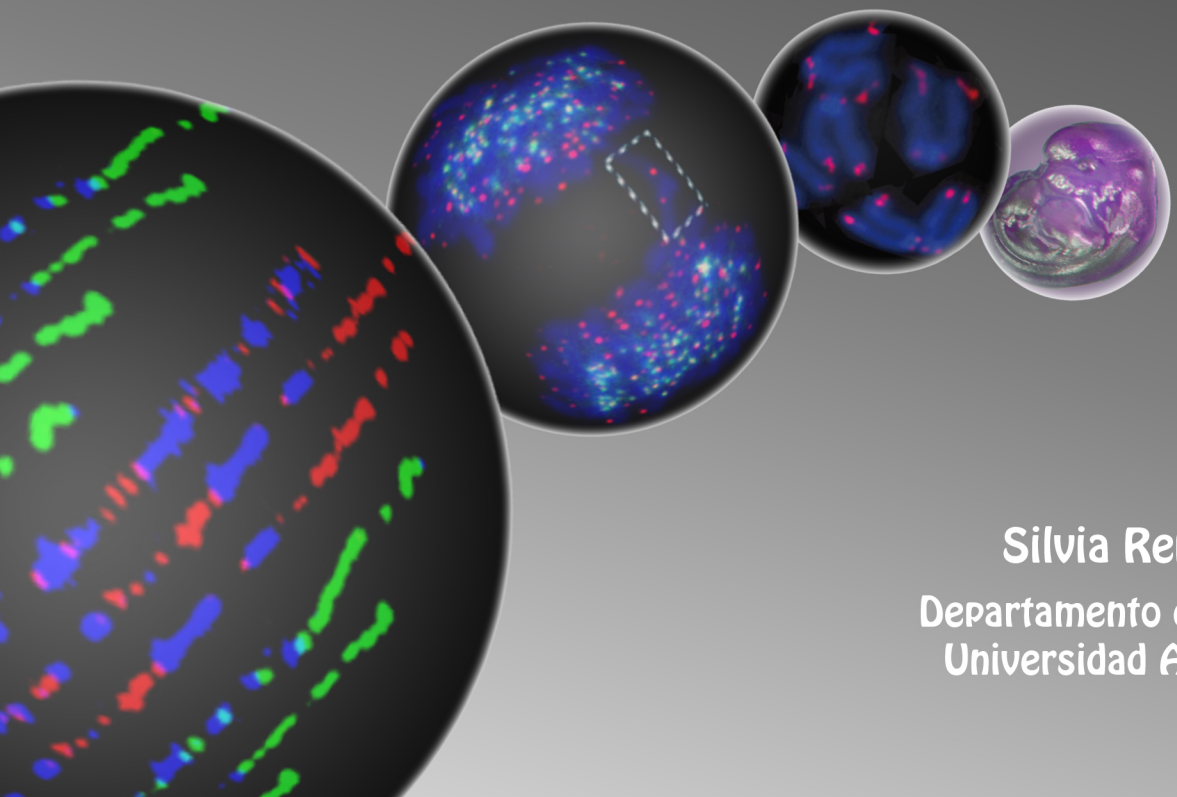


The specific roles of cohesin-SA1 in telomere cohesion and gene expression: Implications for cancer and development



Silvia Remeseiro López

Departamento de Biología Molecular
Universidad Autónoma de Madrid

Universidad Autónoma de Madrid

Facultad de Ciencias

Departamento de Biología Molecular

The specific roles of cohesin-SA1 in telomere cohesion and gene expression:

Implications for cancer and development

Silvia Remeseiro López

Degree in Biology

Thesis Director:

Dr. Ana Losada

Centro Nacional de Investigaciones Oncológicas (CNIO)

Madrid

Ana Losada, Jefa de Grupo del laboratorio de Dinámica Cromosómica del Centro Nacional de Investigaciones Oncológicas

CERTIFICA:

Que la tesis doctoral titulada “The specific roles of cohesin-SA1 in telomere cohesion and gene expression: Implications for cancer and development”, ha sido realizada en el Centro Nacional de Investigaciones Oncológicas y tutelada en el Departamento de Biología Molecular de la Universidad Autónoma de Madrid.

La tesis realizada por Silvia Remeseiro López reúne todas las condiciones requeridas por la legislación vigente y la originalidad y calidad científica para poder ser presentada y defendida con el fin de optar al grado de Doctor.

Y para que conste donde proceda, firmo el presente certificado.

Madrid, a 31 de Julio de 2012

Dra. Ana Losada

A mis padres, a mis hermanas y a Manu

Hay muchas frases que ilustran que el final del camino es el comienzo de uno nuevo. Ha llegado el momento de emprender un nuevo rumbo, no sin temores, pero siempre con muchísimas ganas y con la misma ilusión del primer día. Y no sin antes mostrarle mi cariño a las personas que han caminado junto a mí en los últimos años.

A mi mentora, Ana Losada, que me abrió las puertas de su laboratorio, me acogió con cariño desde el principio, confió en mí y nunca me ha dicho que no a nada. Por todo lo que he aprendido en estos años. Porque, aunque ya lo había dicho Voltaire hace unos cuantos siglos, tú eres quien me ha enseñado que "la ignorancia afirma o niega rotundamente, la ciencia duda". Eres un referente para mí, en la ciencia y en la vida. Mil gracias!

A Ana Cuadrado, por todo lo que hemos aprendido juntas, por lo mucho que hemos disfrutado trabajando y filosofando, por nuestras discusiones sobre ciencia, economía, psicología infantil… o cualquier otra cosa! Contigo he compartido mis mejores momentos en el laboratorio, eres una fuente de energía, optimismo y vitalidad inagotable! Gracias por contagiarme tu entusiasmo, en lo profesional y lo personal. Muchas gracias por Luca, es maravilloso verlo crecer. Pero sobre todo, gracias por tu amistad!

A mis compañeros del grupo de Dinámica Cromosómica, los vecinos de Replicación, los "CDCs" y los "Capetillo", la gente de Telómeros, los bioinformáticos, animalario y servicios internos, a todas las personas del CNIO con las que he trabajado en algún momento. Gracias por vuestra ayuda.

A la gente del laboratorio de Schildkraut en New York, gracias por enseñarme la lección más importante sobre SMARD: tener buenas manos y que los planetas se alineen para que todo salga bien a la primera. I got it!

Oskar, eres un gran referente, muchas gracias por un buen consejo a tiempo!

A mis amigas de siempre, Rosa y Justyna, gracias por seguir ahí a pesar de la distancia.

Y a los grandes amigos que he hecho en esta etapa. Andrés, eres una de las mejores personas que conozco, me siento afortunada de tenerte como amigo. Te echaré de menos, aunque como ya te hemos adoptado, te podías venir de postdoc con nosotros! Chebi y Mery, sois una pareja genial, os quiero; Mery, muchas gracias por escucharme tantas veces y por tu amistad sincera. Pauli, no sabes cuánto te he echado en falta; tú y Ralph sois estupendos, gracias por tantos momentos fantásticos, para cuándo otro roadtrip? A Marianna, Mauro y Samuel, por tantos instantes emotivos y llenos de amor. A todos vosotros, tendréis mi cariño y mi amistad siempre!

Para mis padres nunca tendré suficientes palabras de agradecimiento… Gracias por dejarme elegir mi camino, aun sabiendo que me tropezaría con cientos de piedras. Gracias por dejarme aprender a mí solita, pero estando siempre cerca para tenderme vuestra mano. Os adoro y os echo de menos!

A Bea y Mónica, porque es increíble y genial que siendo hermanas seamos tan diferentes. Gracias por todo lo que me habéis enseñado, a pesar de ser yo la hermana mayor… Ojalá pudiera pasar más tiempo con vosotras. Os quiero!

Y muchas gracias a la vida! A las decisiones que he tomado y me han traído hasta aquí, porque gracias a ello he conocido a la persona que ha iluminado mi camino y expandido mi universo. Manu, gracias por dibujarme una sonrisa en el rostro, siempre, como sólo tú sabes hacer. Se abre un nuevo camino también para nosotros, y sé que será incluso más maravilloso. Con todo mi amor!

Silvia

Contents

Contents

List of abbreviations

3C	Chromosome Conformation Capture
ACA	Anti-Centromeric Antibody
APC/C	Anaphase Promoting Complex/Cyclosome
ATP	Adenosine triphosphate
Cdk1	Cyclin-dependent kinase 1
CdLS	Cornelia de Lange Syndrome
ChIP-seq	Chromatin Immunoprecipitation - sequencing
CoAT	Cohesin Acetyl Transferase
CoDAC	Cohesin Deacetylase
CTCF	CCCTC-binding factor
DNA	Deoxyribonucleic Acid
DSB	Double Strand Break
Eco1	Establishment of cohesion 1, acetyltransferase
ER	Estrogen Receptor
Esco1/2	Establishment of cohesion 1 homolog 1/2
HDAC8	Histone Deacetylase 8
Hos1	Histone deacetylase Hos1
HR	Homologous recombination
IFNG	Interferon Gamma
IGF2-H19	Insulin-like Growth Factor 2 – H19
Igh	Immunoglobulin heavy locus
MEFs	Mouse Embryonic Fibroblasts
Nipbl	Nipped-B homolog (mammalian)
Pds5	Precocious dissociation of sisters
PP2A	Protein Phosphatase 2A
Rad21	Scc1-homolog
RBS	Roberts Syndrome
RNA-seq	RNA-sequencing
RNA	Ribonucleic Acid
SA1/2	Stromal Antigen 1/2
Scc	Sister chromatid cohesion
Sgo	Shugoshin
SMC	Structural Maintenance of Chromosomes
TFs	Transcription Factors
TRF1	Telomeric Repeat binding Factor 1
Wapl	Wings apart-like

Table 1. Cohesin subunits and its regulatory factors.

	S. cerevisiae	*S. pombe*	*D. melanogaster*	*X. laevis*	*H. sapiens / M. musculus*
SMC	Smc1	Psm1	Smc1	Smc1	Smc1a, Smc1b
	Smc3	Psm3	Smc3	Smc3	Smc3
a-Kleisin	Scc1/Mcd1	Rad21	Rad21	Rad21	Scc1/Rad21
	Rec8	Rec8	C(2)M	-	Rad21L, Rec8
a-Kleisin interacting subunits	Scc3	Psc3	SA	SA1, SA2	SA1/Stag1, SA2/Stag2
	-	Rec11	-	-	SA3/Stag3
Regulatory factors	Pds5	Pds5	Pds5	Pds5A, Pds5B	Pds5A, Pds5B/APRIN
	Rad61/Wapl	Wapl	Wapl	Wapl	Wapl/Wapal
	-	-	Dalmatian	Sororin	Sororin
Cohesin loading complex	Scc2	Mis4	Nipped-B	Scc2	Nipbl/Scc2
	Scc4	Ssl3	Scc4	Scc4	Mau2/Scc4
Cohesin Acetyl Transferases	Eco1/Ctf7	Eso1	Deco, San	Esco1, Esco2	Esco1, Esco2
Cohesin Deacetylases	Hos1	-	-	-	HDAC8

Proteins in red correspond to meiotic isoforms.

Summary

Cohesin is a protein complex originally identified for its role in sister chromatid cohesion. Increasing evidence portrays this complex as a major organizer of interphase chromatin. In vertebrate somatic cells, there are two cohesin complexes that consist of Smc1, Smc3, Rad21 and either SA1 or SA2. To explore their functional specificity and their relevance for cancer and development, we generated a mouse model deficient for SA1. Mouse embryos lacking SA1 die before birth and show developmental delay and features reminiscent of Cornelia de Lange syndrome (CdLS), a genetic disorder linked to cohesin dysfunction. SA1-heterozygous mice have shorter lifespan and earlier onset of tumourigenesis. SA1-deficient cells present decreased proliferation and increased aneuploidy as a result of chromosome segregation defects. These defects are not caused by impaired centromere cohesion, which depends on cohesin-SA2. Instead, we found that they arise from defective telomere replication, which requires cohesion mediated specifically by cohesin-SA1. In addition, analysis of cohesin genome-wide distribution reveals that SA1 is responsible for the accumulation of the complex at promoters and sites bound by the chromatin insulator CTCF. In the absence of SA1, cohesin-SA2 redistributes away from those sites and transcription of several genes is altered. Among them we found genes involved in biological processes related to CdLS pathogenesis. The presence of cohesin-SA1 at gene promoters positively regulates the expression of genes such as Myc or Protocadherins, a function that cannot be compensated by cohesin-SA2. Moreover, the cohesin-binding pattern along gene clusters is altered in the absence of SA1 and affects the transcriptional regulation of the genes within.

Thus, cohesin-SA1 prevents generation of aneuploidy and tumourigenesis by assuring an efficient telomere replication, while its impaired function in gene expression regulation most likely underlies the etiology of CdLS.

Resumen

El complejo cohesina se identificó inicialmente por su papel en cohesión, pero evidencias más recientes lo describen como organizador de la cromatina interfásica. En células somáticas de vertebrados hay dos tipos de complejos compuestos por Smc1, Smc3, Rad21 y bien SA1 o SA2. Para estudiar su especificidad funcional y relevancia en cáncer y desarrollo embrionario, generamos un modelo murino deficiente en SA1. Los embriones deficientes en SA1 no son viables, presentan un retraso en el desarrollo y características propias del síndrome de Cornelia de Lange (CdLS), asociado a una disfunción en cohesina. Los ratones SA1 heterocigotos presentan reducción en la esperanza de vida y aparición temprana de tumores espontáneos. Las células deficientes en SA1 muestran una menor capacidad proliferativa y mayor aneuploidía, como resultado de los defectos en segregación cromosómica. Estos últimos no se deben a defectos en cohesión centromérica, mediada por SA2, sino que tienen su origen en una replicación telomérica deficiente, al faltar la cohesión mediada específicamente por SA1 en esta región. Además, el análisis de la distribución genómica de cohesina revela que SA1 es responsable de la acumulación del complejo en promotores y sitios de unión del factor *insulator* CTCF. En ausencia de SA1, la cohesina-SA2 se desplaza a otras localizaciones y se altera la transcripción de varios genes. Entre ellos se encuentran genes implicados en procesos biológicos relacionados con la patogénesis de CdLS. La presencia de SA1 en promotores regula positivamente la expresión de genes como Myc y Protocaderinas, función que no puede ser compensada por SA2. Además, las posiciones de cohesina en *clústeres* génicos se alteran en ausencia de SA1, lo cual afecta a su regulación transcripcional.

La cohesina-SA1 previene la generación de aneuploidía y tumores al garantizar una replicación telomérica eficiente. Asimismo, es probable que su papel en regulación de la expresión génica guarde relación directa con la etiología de CdLS.

“Ignorance affirms or denies wholeheartedly. Science doubts”.

Voltaire

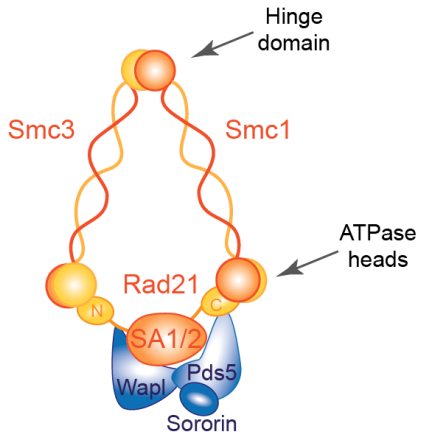
Introduction

Cohesin is a protein complex originally identified for its role in sister chromatid cohesion (Guacci et al. 1997; Michaelis et al. 1997; Losada et al. 1998), although it has been recently portrayed also as a major organizer of interphase chromatin. Its cohesive function is essential for accurate chromosome segregation and for homologous recombination (HR)-mediated repair in G2 (Nasmyth and Haering 2009). Cohesin acting as chromatin loop organizer was reported to contribute to gene expression regulation of certain loci (Hadjur et al. 2009; Mishiro et al. 2009; Nativio et al. 2009; Kagey et al. 2010; Seitan et al. 2011), spatial organization of DNA replication factories (Guillou et al. 2010) and recombination (Degner et al. 2009; Seitan et al. 2011). Despite being discovered years ago, many aspects of cohesin complex functions remain poorly understood. In particular, the specific role of different cohesin complexes present in somatic cells, that contain either SA1 or SA2 subunit, was unknown until recently.

Cohesin complex: composition, architecture and mode of action

Cohesin is one of the three SMC (Structural Maintenance of Chromosomes) complexes present in eukaryotes, the other two being condensin and the Smc5/6 complex. It is composed of Smc1 and Smc3, the kleisin subunit Rad21/Scc1 and the SA/Scc3 subunit (Figure 1; see Table1 for nomenclature on different organisms on page 16).

Figure 1

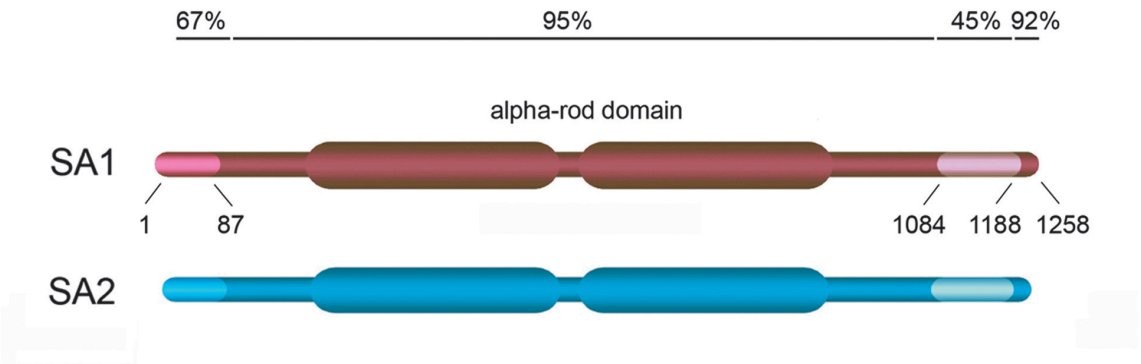


Schematic representation of the cohesin complex and its regulatory factors. Cohesin subunits Smc1, Smc3, Rad21 and either SA1 or SA2 are depicted in orange, while the blue figures represent the regulatory factors Pds5, Wapl and Sororin, that modulate the chromatin-association of the complex.

SMC proteins are long polypeptides that fold back on themselves by anti-parallel coiled-coil interactions forming a heterodimer molecule with a "hinge" domain at one end and a globular ATPase "head" at the other. The hinge domains of Smc1 and Smc3 are tightly bound, whereas the ATPase heads are connected by Rad21 forming a ring-shaped molecule. Rad21 is additionally connected with a fourth subunit, SA, which exists in two different versions in somatic vertebrate cells named stromal antigen 1 and 2 (SA1 and SA2). These two SA proteins never coexist in the same complex. Thus, there are two different versions of cohesin that contain either one or the other and that we refer to as cohesin-SA1 and cohesin-SA2.

SA proteins are around 1250 amino acid long and present more than 95% homology along their central region, which contains a-rod repeats (Palidwor et al. 2009). The N- and C-terminal regions are more divergent (Figure 2). Cohesin-SA2 is 3-10 times more abundant than cohesin-SA1 in human and *Xenopus* somatic cells, while *Xenopus* eggs contain 10 times more cohesin-SA1 than cohesin-SA2 (Losada et al. 2000; Holzmann et al. 2011). In addition, immunofluorescent staining indicates that both SA proteins distribute throughout interphase chromatin in an indistinguishable manner and present similar regulation throughout the cell cycle (Losada et al. 2000; Sumara et al. 2000).

Figure 2



Representation of SA proteins domains and their regions of maximal divergence. The percentages refer to the degree of homology between both proteins in each of the depicted regions (N-terminus, central alpha-rod and C-terminus domains), whose length is indicated by the residue number that they comprise.

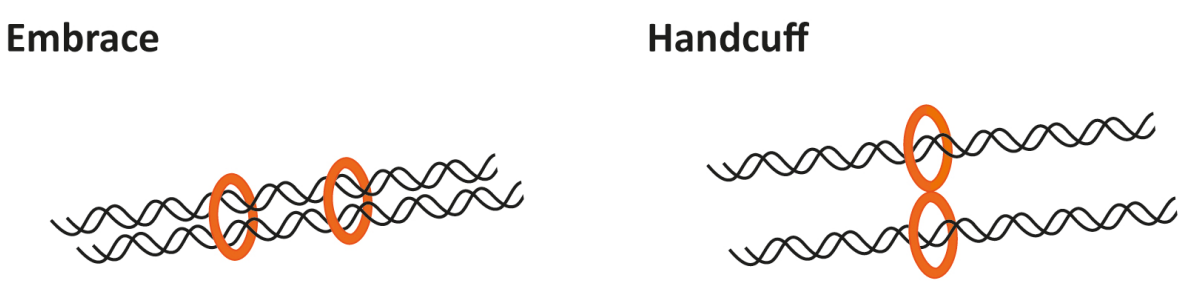
Based on the high degree of homology between the two SA proteins and their similar behavior in terms of chromatin association during the cell cycle, it has been largely assumed that SA1 and SA2 have redundant functions, or even that SA2 is more relevant because it is more abundant. However, SA1 or SA2 downregulation in HeLa cells by RNA interference has suggested that cohesin-SA2 is specifically required for centromeric cohesion whereas cohesin-SA1 is responsible for arm and telomere cohesion (Canudas and Smith 2009). Intriguingly, downregulation of a telomere protein named TIN2, part of the shelterin complex, produced similar defects in both telomere and arm cohesion. To us, it was not clear why a telomere protein would have the ability to regulate arm cohesion.

On other hand, cohesin behavior is modulated by close interaction with other regulatory factors: Pds5, Wapl and Sororin, which associate with chromatin through the cohesin complex (Figure 1). Pds5 and Wapl work together to modulate the association of cohesin with chromatin. They interact with each other and their binding to Rad21 is promoted or

stabilized by the SA subunit. Vertebrate cells have two Pds5 proteins, Pds5A and Pds5B, and both bind Wapl as well as cohesin-SA1 and cohesin-SA2 (Kueng et al. 2006; Losada 2005; Gandhi et al. 2006), but they are never together in the same complex. The third regulatory factor, Sororin, is essential for cohesion establishment and maintenance (Schmitz et al. 2007). One of the major roles of Sororin is to antagonize Wapl to maintain sister chromatid cohesion (Nishiyama et al. 2010).

Given the ring shape of the complex, cohesin is able to embrace the DNA fiber and, therefore, it acts as topological linker. There are different models for the mechanism by which cohesin mediates sister chromatid cohesion (Figure 3). The embrace model proposes that a single complex entraps both sister chromatids (Haering et al. 2008). The less popular handcuff model postulates that cohesion would be performed by the association of two cohesin complexes entrapping a single chromatid each, maybe through their SA subunits (Zhang et al. 2008b) or by additional factors.

Figure 3



Two models to explain how cohesin establishes sister chromatid cohesion: the embrace model (left) and the handcuff model (right). For simplicity, cohesin is represented as an orange ring that entraps the sister chromatids. In the handcuff model, a single SA subunit interacts with two Smc1-Smc3-Rad21 rings.

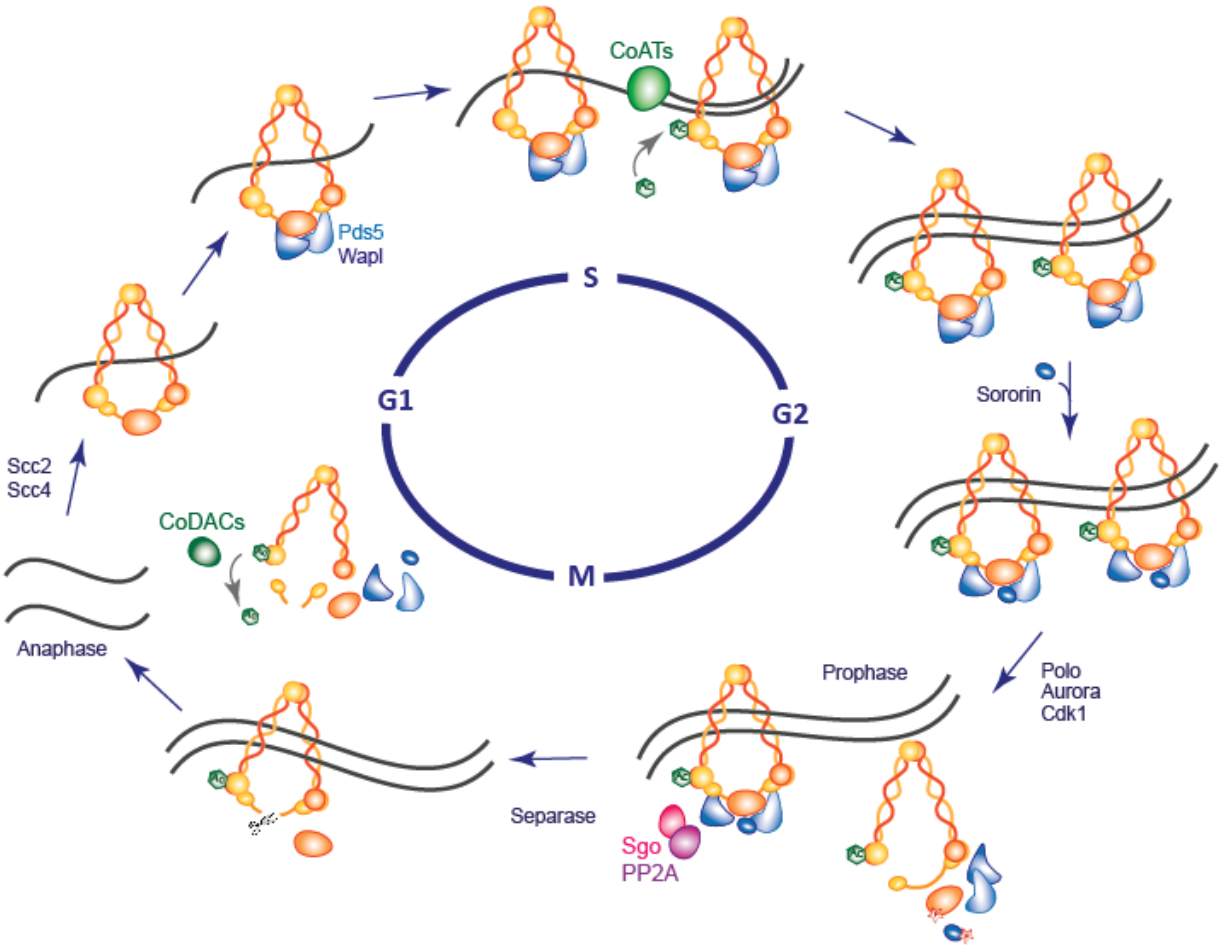
Cohesin regulation through the cell cycle

The association of cohesin with chromatin is tightly regulated through the cell cycle (Figure 4). In vertebrate cells, cohesin is loaded onto chromatin in early G1. This loading depends on the heterodimer complex formed by Nipbl/Scc2 and Scc4 (Seitan et al. 2006; Watrin et al. 2006). It is not known how Nipbl is recruited to chromatin for cohesin loading or how the loading sites are specified. Some evidences point out that Nipbl-Scc4 promote cohesin entrapment of the DNA fiber by favoring the ATP hydrolysis by the SMC heads, which might require dissociation of the Smc1/3 hinge domains (Milutinovich et al. 2007; Gruber et al. 2006). How ATP hydrolysis at the SMC heads drives hinge opening is not understood but likely involves the Scc3/SA subunit (Hu et al. 2011; Kurze et al. 2011).

The establishment of cohesion in S phase requires the acetylation of the Smc3 subunit by cohesin acetyltransferases (CoATs, e.g. Eco1, Esco1/2), that acetylate two lysine residues located in the Smc3 head domain (Unal et al. 2008; Rolef Ben-Shahar et al. 2008; Zhang et al. 2008a). This modification counteracts the anti-establishment function of Wapl-Pds5 (that prevents cohesin to become cohesive) most likely by destabilizing cohesin binding to Wapl-Pds5 and, thereby, switching the conformation to allow cohesion establishment (Terret et al. 2009). In yeast, Smc3 acetylation by Eco1 prevents the dissociation of the Smc3/a-kleisin interface, recently reported to be the "cohesin's DNA exit gate" (Chan et al. 2012). In vertebrates, Smc3 acetylation is not sufficient for cohesion establishment, maybe because Wapl-Pds5 antiestablishment is stronger. Acetylation-dependent binding of Sororin is essential to stabilize a fraction of cohesin after DNA replication, antagonizing Wapl to maintain sister chromatid cohesion: Pds5 provides a binding surface to Sororin, which would in this way displace Wapl (Schmitz et al. 2007; Nishiyama et al. 2010; Lafont et

al. 2010). Importantly, all cohesin subunits are sumoylated during S phase in budding yeast and non-sumoylated cohesin complexes can be acetylated but are not cohesive (Almedawar et al. 2012). How sumoylation cooperates with acetylation and the conservation of this modification in other eukaryotes has not been reported yet.

Figure 4



Regulation of cohesin-chromatin association throughout the cell cycle in vertebrate cells.

Cohesion is maintained during G2, when cohesin plays a role in promoting DNA repair by homologous recombination (Sjögren and Nasmyth 2001). In yeast, cohesin is loaded *de novo* by Scc2-Scc4 in a large region surrounding a double strand break (DSB) and also

genome-wide. Moreover, cohesion is established upon DNA damage in G2 cells in an Eco1 dependent manner, although in this case Scc1 is acetylated (Sjögren and Ström 2010). Scc1 is also sumoylated by Mms21/Nse2, a subunit of the Smc5/6 complex (McAleenan et al. 2012). Thus, the global reinforcement of cohesion upon occurrence of a DSB may help prevent promiscuous pairing of the free end, in particular between homologous chromosomes in diploid cells (Covo et al. 2010). In addition, interaction between BRCA2 and cohesin through Pds5 is important for effective homologous recombination (HR)-mediated DNA repair and normal response to DNA damage in human cells (Brough et al. 2012). These observations support the relevance of cohesin in promoting faithful repair in higher eukaryotes, as in yeast (Heidinger-Pauli et al. 2010). Cohesin also facilitates the recruitment of proteins involved in the activation of the intra-S and G2/M checkpoints (Watrin and Peters 2009).

Sister chromatid cohesion is maintained until mitosis, when it must be resolved to allow the separation of the genetic material into the two daughter cells (Figure 4). In metazoan, the complex dissociates from chromatin in two steps in prophase and anaphase. More than 90% of cohesin is released in the prophase pathway upon phosphorylation of the SA subunit by Polo and the action of Pds5-Wapl (Shintomi and Hirano 2010), which is favored by the removal of Sororin after Cdk1 phosphorylation (Nishiyama et al. 2010; Dreier et al. 2011). The remaining cohesin, enriched at pericentromeric regions, is protected from the prophase pathway by Shugoshin 1 (Sgo1) and the protein phosphatase PP2A by counteracting SA phosphorylation (Gutiérrez-Caballero et al. 2012). This cohesin population ensures chromatid cohesion until all chromosomes achieve biorientation. At the onset of anaphase, activation of the APC/C targets securin and cyclin B for degradation, initiating mitotic exit. At this step, complete dissolution of cohesion requires separase-mediated cleavage of Rad21/Scc1 and the decatenating activity of topoisomerase II (Oliveira et al.

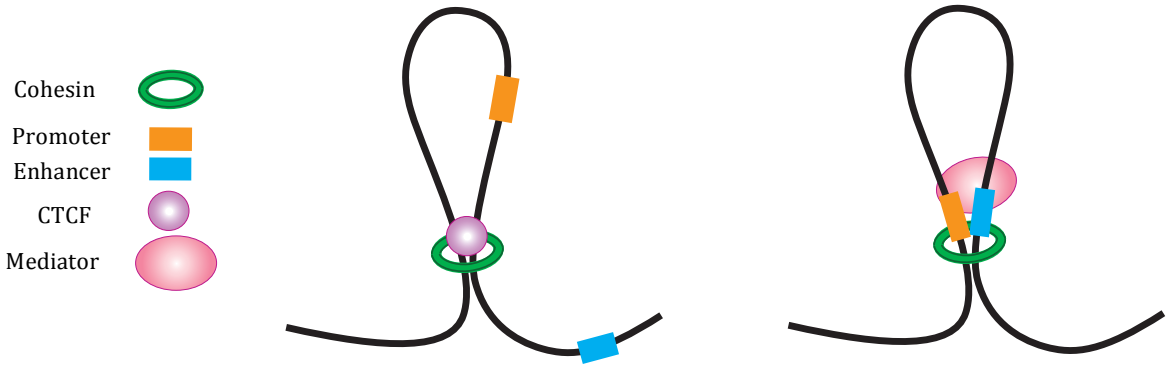
2010; Shamu and Murray 1992; Wang et al. 2008). In order to be reused in the next cycle, all cohesin molecules must be deacetylated by cohesin deacetylases (CoDACs) Hos1 in yeast (Borges et al. 2010; Beckouët et al. 2010) and HDAC8 in human cells (Deardorff et al. 2012a).

Cohesin as an organizer of the interphase chromatin

Cohesin is loaded on chromatin long before cohesion is established in vertebrate cells and in high excess in relation to what is needed to maintain cohesion in mitosis. Moreover, cohesin is present in non-proliferating cells (Wendt et al. 2008) and inside transcriptional units in different organisms (Wendt et al. 2008; Parelho et al. 2008; Misulovin et al. 2008). Together with previous results in Drosophila, these data pointed to a role of cohesin in regulation of gene expression (Dorsett 2011). With the finding of mutations in the components of cohesin pathway as the cause of Cornelia de Lange Syndrome (see section below), the role of the complex in transcriptional regulation emerged as a hypothetical mechanism underlying the human syndrome.

A key discovery supporting a role of cohesin in transcription was the finding that many cohesin-binding sites, identified by Chromatin Immunoprecipitation (ChIP) techniques in mouse and human cells, colocalize with sites also bound by the insulator CTCF, thus mediating transcriptional insulation (Wendt et al. 2008; Parelho et al. 2008; Rubio et al. 2008) (Figure 5, left). However, many aspects of cohesin function in transcription remain unclear, and it seems that the transcriptional machinery deals with cohesin bound to chromatin in different ways depending on the organisms.

Figure 5



Schematic representation of some mechanisms by which cohesin performs gene expression regulation. (Left) Cohesin mediates transcriptional insulation by colocalization with CTCF at certain genomic positions. (Right) Cohesin and mediator determine promoter-enhancer interactions.

In yeast, a negative association between transcription and cohesin binding has been shown, since cohesin localizes almost exclusively between genes that are transcribed in opposing directions (Lengronne et al. 2004; Glynn et al. 2004). Cohesin is loaded by Scc2-Scc4 at certain sites from where it seems to slide towards its permanent positions, being excluded from actively transcribed genes and accumulated between genes of convergent transcription. In *S. pombe*, cohesin is not only accumulated between convergent genes, but it also acts as a terminator of transcription (Gullerova and Proudfoot 2008). In contrast, in *Drosophila* cohesin binds preferentially to active genes where it colocalizes with Nipped-B/Scc2 and RNA polymerase II (Misulovin et al. 2008), and it is absent from repressive regions characterized by binding of Polycomb proteins and H3K27me3 (Schaaf et al. 2009). In *Drosophila* salivary glands cohesin directly regulates the expression of a set of genes, including the ones that mediate the response to ecdysone, by direct binding to most of

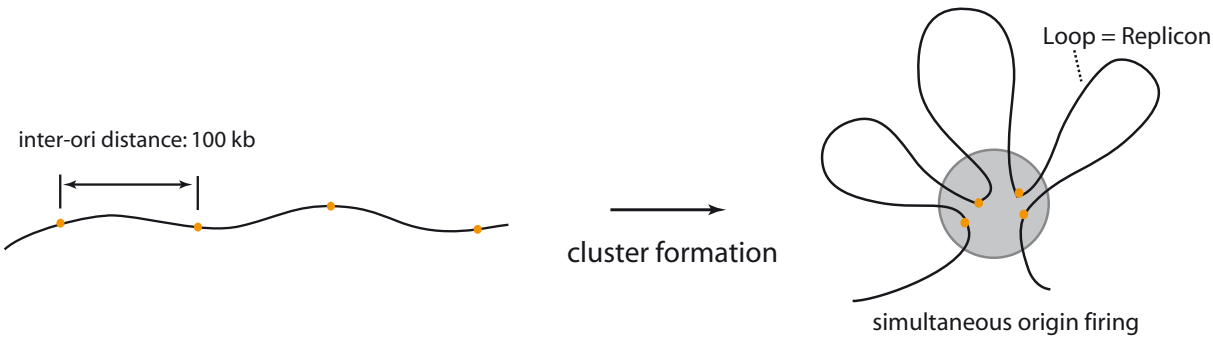
these loci (Pauli et al. 2010). Cohesin also binds and regulates genes with paused RNA polymerase, and most likely hinders transition to transcriptional elongation (Fay et al. 2011).

As mentioned earlier, cohesin largely colocalizes with the transcriptional insulator CTCF in human and mouse cells and mediates transcriptional insulation (Wendt et al. 2008; Parelho et al. 2008; Rubio et al. 2008). Chromosome conformation capture (3C) studies suggest that cohesin performs these transcriptional functions by promoting the formation of chromatin loops in collaboration with factors such as CTCF (Figure 5, left). Given the binding of cohesin to chromatin by topologically embracing the DNA fiber, it is unlikely that cohesin has affinity for a particular DNA sequence. Cohesin might recognize the structure of a loop itself, or instead, it might directly interact with CTCF (Stedman et al. 2008). By embracing the chromatin fiber in *cis* at the base of such loops, cohesin would stabilize them.

There are several evidences suggesting that cohesin forms the topological and mechanistic basis for long-range chromosomal interactions in *cis.* Such are the cases of the developmentally regulated locus IFNG in a cell-type specific manner (Hadjur et al. 2009); the human *apolipoprotein* gene region, whose alterations increase the risk of dyslipidemia and atherosclerosis (Mishiro et al. 2009); the imprinted *IGF2-H19* locus (Nativio et al. 2009) or the beta-globin locus (Chien et al. 2011) and the surrounding olfactory receptor genes (Hou et al. 2010). Cohesin also colocalizes with transcriptional regulators other than CTCF, such as the estrogen receptor-alpha (ER) in breast cancer cells, although a direct interaction between them has not been demonstrated (Schmidt et al. 2010). In murine embryonic stem cells, cohesin and the transcriptional coactivator Mediator facilitate DNA looping between the enhancers and promoters of some critical genes required to maintain pluripotency (Kagey et al. 2010) (Figure 5, right).

Cohesin also contributes to processes other than transcription in the interphase nucleus. Upon DSB induction, cohesin restricts gH2AX accumulation at neighboring genes to preserve their activity (Caron et al. 2012). Cohesin interacts with components of the pre-Replication Complex and participates in the spatial organization of DNA replication factories in human cells by organizing them in chromatin loops (Guillou et al. 2010) (Figure 6). Moreover, cohesin and CTCF are recruited to the immunoglobulin loci during B lymphocyte development (Degner et al. 2009) and participate in V(D)J recombination by influencing the architecture of the Igh locus and the antisense transcription in pro-B cells (Degner et al. 2011). Deletion of Rad21 in a population of non-dividing mouse thymocytes showed that cohesin plays a role in T-cell-receptor a locus rearrangement, as well as in its transcription, by promoting long-range interactions within this megabase-sized locus (Seitan et al. 2011).

Figure 6



Model for the architectural role of cohesin at DNA replication factories. Replication origins (orange dots) within a DNA region are grouped in rosette-like structures by the action of cohesin.

Cohesin and human disease: from cohesinopathies to cancer

The term cohesinopathies was coined to describe syndromes associated with the effect of mutations in the regulators and structural components of the cohesin complex. To date, the two major cohesinopathies studied are Cornelia de Lange Syndrome (CdLS) and Roberts syndrome (RBS)/SC Phocomelia (Tonkin et al. 2004; Krantz et al. 2004; Vega et al. 2005; Musio et al. 2006). CdLS is caused by heterozygous mutations in the cohesin loader Nipbl in approximately half of the cases (Krantz et al. 2004) and, in a reduced frequency, in the cohesin subunits Smc1, Smc3, Rad21 and the CoDAC HDAC8 (Musio et al. 2006; Deardorff et al. 2007; 2012b; 2012a). Roberts syndrome and SC Phocomelia have been linked to homozygous mutations in Esco2 (Vega et al. 2005; Schüle et al. 2005).

CdLS is a dominantly inherited disorder that affects 1 in 10,000-30,000 live births. This syndrome is characterized by mental retardation, reduced body size, dysmorphic face, upper limb defects and several additional organ anomalies (Liu and Krantz 2009) (Figure 7).

Figure 7

Representative images of CdLS patients showing features of the syndrome such as (a) characteristic facial features and upper limb abnormalities including (b) oligodactyly or (c) missing and fused digits. From Krantz et al. 2004.

Cells from CdLS patients do not show overt cohesion defects but instead, microarray data reveals altered transcriptional profiles (Castronovo et al. 2009; Liu et al. 2009). Altered transcription is also recapitulated in Nipbl-deficient mouse and zebrafish models of the syndrome (Kawauchi et al. 2009; Muto et al. 2011). In contrast, RBS is a rare autosomal recessive disorder, characterized also by mental retardation and growth deficiency, but involving loss of cohesion. This is consistent with this syndrome being associated to mutations in the acetyltransferase that promotes cohesion establishment and, therefore, cells derived from either patients or zebrafish model do present cohesion defects (Vega et al. 2005; Mönnich et al. 2011).

On other hand, and given the essential role of cohesin for accurate chromosome segregation, it was reasonable to hypothesize that cohesion defects might generate aneuploidy. These aberrations in chromosome number are a feature commonly observed in cells from different malignant tumors and have been proposed to contribute to tumourigenesis. However, until recently, little was known of the involvement of cohesin mutations in human cancer (Xu et al. 2011). Mutations in cohesion factors were initially identified in human colorectal cancer (Barber et al. 2008). More recently, inactivating mutations in SA2 have been described in a diverse range of human tumor types (Solomon et al. 2011). This report suggests that lack of SA2 function leads to aneuploidy due to precocious dissociation of sister chromatids in metaphase, in agreement with previous reports pointing to the role of SA2 in centromeric cohesion (Canudas and Smith 2009). Cohesin mutations were also recently found in acute myeloid leukemia patients (Welch et al. 2012). However, many questions remain open in the field (Xu et al. 2011). What roles cohesin plays in the response to cancer therapy/prognosis or what precise cohesin-mediated mechanisms underlie initiation and progression of different cancer types are issues that remain to be studied in the future.

Mouse models to study cohesin functions *in vivo*

The relevance of cohesin functions in cell proliferation and differentiation has been increasing rapidly (Merkenschlager 2010; Dorsett 2011). Mouse models are powerful tools to unravel these different functions of cohesin, together with their temporal and spatial regulation; e.g. during embryo development/adult age and in different tissues, respectively. To date, mice deficient for genes encoding cohesin components have been generated for meiosis-specific subunits (Smc1b, Rec8 and Rad21L) as well as for Rad21. The cohesin subunit Smc1b plays a key role in meiotic cohesion, DNA recombination, synapsis and chromosome movements in meiosis (Revenkova et al. 2004). Smc1b-deficient mice (Revenkova et al. 2004) as well as Rec8-null mice (Xu et al. 2005) of both sexes are sterile, although because of different anomalies in meiosis. It has been also reported that Rad21L, a recently identified kleisin, is required for full synapsis of homologous chromosomes and its absence provokes male sterility as a consequence of total azoospermia, whereas females are fertile but develop infertility as they age (Herrán et al. 2011). Biallelic inactivation of Rad21 results in early embryonic lethality and only the radiosensitivity of Rad21 heterozygous animals was reported (Xu et al. 2010). A mouse with a conditional allele of Rad21 has also been generated and employed for depleting cohesin specifically in non-cycling thymocytes (Seitan et al. 2011). Knockout mouse models for the cohesin regulatory factors Pds5A and Pds5B die soon after birth and present developmental defects similar to those observed in CdLS patients (Zhang et al. 2007; 2009). The effects of the combined deficiency of both genes suggest that Pds5A and Pds5B have some overlapping functions and a minimum dosage is required for proper development. Importantly, heterozygous mice for the gene encoding the cohesin loader Nipbl, that is mutated in around 50% of CdLS cases, show many features characteristic of the human syndrome

(Kawauchi et al. 2009). Complete ablation of Esco2, that is mutated in RBS patients, results in early embryonic lethality In mice due severe impairment of cohesion and chromosome segregation (Whelan et al. 2012).

The work presented in this thesis is the first thorough characterization of a mouse model deficient for a ubiquitously expressed cohesin subunit, SA1, and its implications for such relevant human diseases as cancer and CdLS. We aimed to generate and characterize a knockout mouse model for cohesin subunit SA1 in order to dissect the functional specificity of the two different cohesin complexes, cohesin-SA1 and cohesin-SA2, in cellular processes such as cohesion, chromosome segregation and transcriptional regulation. We also wished to assess the effect of SA1 ablation on embryonic development and its connection with human disease, in particular cancer and CdLS. We found that SA1-deficiency results in delayed embryo development and lethality, which shows that cohesin-SA1 and cohesin-SA2 are not redundant. Defects in cohesin-SA1 function generate aneuploidy and result in increased incidence of tumourigenesis due to impaired replication of telomeres. Moreover, we demonstrate that cohesin-SA1 is responsible for cohesin distribution and accumulation at promoters and CTCF sites. In the absence of SA1, cohesin becomes redistributed, which alters the expression of genes involved in developmental pathways relevant to CdLS.

“Imagination is more
important than knowledge”.
Albert Einstein

Aim of the work

1. Generate a knockout mouse model for the cohesin subunit SA1.

2. Study the effect of cohesin-SA1 deficiency during embryonic development and in human diseases such as cancer and CdLS.

3. Dissect the functional specificity of cohesin-SA1 and cohesin-SA2 in terms of sister chromatid cohesion and chromosome segregation.

4. Define a complete map of the genome-wide cohesin binding sites for cohesin-SA1 and cohesin-SA2.

5. Assess the role of cohesin-SA1 and cohesin-SA2 in transcriptional regulation.

“Science is
the poetry of reality”.
Richard Dawkins

Results

Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada.

The EMBO Journal (2012) 31, 2076–2089

The work presented here represents the first thorough characterization of a mouse model deficient for a ubiquitously expressed cohesin subunit, SA1. We have generated this system, which ensures complete absence of the protein, to gain insight into the functional specificity of SA1 and SA2 containing cohesin complexes. We have transcended the cellular level in our studies and addressed the relevance of SA1 functions in embryonic development, which is particularly important given the existence of a human developmental disorder known as Cornelia de Lange syndrome that is linked to cohesin dysfunction. We have shown that SA1 deficiency results in delayed embryo development and lethality, demonstrating that the functions of both SA proteins are not redundant. For the first time, we report that defects in cohesin-SA1 function generate aneuploidy and result in increased incidence and earlier onset of tumourigenesis. Unlike cohesin-SA2, lack of cohesin-SA1 leads to aneuploidy in the absence of precocious separation of sister chromatids in metaphase. Instead, chromosome segregation defects are caused by defective telomere replication in the absence of telomere cohesion, which is mediated exclusively by cohesin-SA1. Telomeres are regions difficult to replicate and forks stall with high frequency. Cohesin prevents collapse of these forks and/or helps their restart. We speculate that unreplicated telomeres prevent proper chromosome segregation of sister chromatids during anaphase, causing lagging chromosomes and chromatin bridges in SA1-deficient cells. Thus, we have proposed a new mechanism for aneuploidy generation due to impaired telomere replication.

My contribution to this study comprises the design, performance, analysis and interpretation of all the experiments. I also participated in the writing of the manuscript, with input from the rest of the authors and under the supervision of the thesis director Dr. Ana Losada.

The EMBO Journal (2012) 31, 2076–2089 | © 2012 European Molecular Biology Organization | All Rights Reserved 0261-4189/12
www.embojournal.org

THE EMBO JOURNAL

Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres

Silvia Remeseiro[1], Ana Cuadrado[1], María Carretero[1], Paula Martínez[2], William C Drosopoulos[3], Marta Cañamero[4], Carl L Schildkraut[3], María A Blasco[2] and Ana Losada[1,*]

[1]Chromosome Dynamics Group, Molecular Oncology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain, [2]Telomeres and Telomerase Group, Molecular Oncology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain, [3]Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY, USA and [4]Comparative Pathology Unit, Biotechnology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Cohesin is a protein complex originally identified for its role in sister chromatid cohesion, although increasing evidence portrays it also as a major organizer of interphase chromatin. Vertebrate cohesin consists of Smc1, Smc3, Rad21/Scc1 and either stromal antigen 1 (SA1) or SA2. To explore the functional specificity of these two versions of cohesin and their relevance for embryonic development and cancer, we generated a mouse model deficient for SA1. Complete ablation of SA1 results in embryonic lethality, while heterozygous animals have shorter lifespan and earlier onset of tumourigenesis. SA1-null mouse embryonic fibroblasts show decreased proliferation and increased aneuploidy as a result of chromosome segregation defects. These defects are not caused by impaired centromeric cohesion, which depends on cohesin-SA2. Instead, they arise from defective telomere replication, which requires cohesion mediated specifically by cohesin-SA1. We propose a novel mechanism for aneuploidy generation that involves impaired telomere replication upon loss of cohesin-SA1, with clear implications in tumourigenesis.

The EMBO Journal (2012) **31,** 2076–2089. doi:10.1038/emboj.2012.11; Published online 13 March 2012
Subject Categories: cell cycle; genome stability & dynamics
Keywords: cancer; chromosome segregation; cohesion; embryonic development; mouse model

Introduction

Cohesin was originally identified as a mediator of sister chromatid cohesion (Guacci *et al*, 1997; Michaelis *et al*, 1997; Losada *et al*, 1998). Its cohesive function is essential for accurate chromosome segregation and for homologous recombination (HR)-mediated DNA repair in G2 (Nasmyth and Haering, 2009). More recent evidence suggests that cohesin acts as a chromatin loop organizer in the interphase nucleus and thereby contributes to regulate gene expression, at least of certain loci (Hadjur *et al*, 2009; Mishiro *et al*, 2009; Nativio *et al*, 2009; Kagey *et al*, 2010; Seitan *et al*, 2011). It also participates in the spatial organization of DNA replication factories, which in turn determines origin activation and thus the efficiency of the duplication process (Guillou *et al*, 2010). Cohesin is composed of a heterodimer of Smc1 and Smc3, the kleisin subunit Rad21 (also known as Scc1/Mcd1) and a protein called Scc3 in yeast and stromal antigen (SA) in vertebrates, that form a ring-like structure that embraces the chromatin fibre(s) (Anderson *et al*, 2002; Haering *et al*, 2008). Two additional proteins named Pds5 and Wapl interact closely with cohesin to modulate its binding to chromatin and are sometimes considered cohesin subunits (Losada *et al*, 2005; Gandhi *et al*, 2006; Kueng *et al*, 2006; Gause *et al*, 2010). Another cohesion factor named Sororin associates with and stabilizes the cohesin rings embracing two sister chromatids after DNA replication (Lafont *et al*, 2010; Nishiyama *et al*, 2010).

In somatic vertebrate cells, there are two versions of the Scc3/SA subunit, SA1 and SA2 (Carramolino *et al*, 1997; Losada *et al*, 2000; Sumara *et al*, 2000). They are ~1250 amino acid long proteins with >75% sequence identity along their central core, which contains two long regions of α-rod repeats (Palidwor *et al*, 2009). The N-terminal and C-terminal domains are more divergent. Cohesin-SA2 (i.e., cohesin containing the SA2 subunit) is 3–10 times more abundant than cohesin-SA1 in human (HeLa) and Xenopus somatic cells, whereas Xenopus eggs contain 10 times more cohesin-SA1 than cohesin-SA2 (Losada *et al*, 2000; Holzmann *et al*, 2010). The functional differences between the two complexes are yet to be clarified. Immunofluorescent staining indicates that SA1 and SA2 distribute throughout interphase chromatin in an indistinguishable manner (Losada *et al*, 2000; Sumara *et al*, 2000). Genome-wide analyses using chromatin immunoprecipitation (ChIP) techniques also suggest that there are no significant differences between the distribution of a common cohesin subunit like Rad21 and either SA2 (Wendt *et al*, 2008) or SA1 (Rubio *et al*, 2008). Both cohesins are released in prophase and only a small fraction is left by metaphase to ensure chromosome alignment (Vagnarelli *et al*, 2004; Toyoda and Yanagida, 2006). SA1 and SA2 can be detected at the centromeres of mitotic chromosomes from Xenopus egg extracts and human cells, respectively (Losada *et al*, 2000; Hauf *et al*, 2005) and Sgo1-PP2A, the protector of centromeric cohesin, can reverse the phosphorylation of both SA subunits by Polo, at least *in vitro* (Rivera and Losada, 2009). However, downregulation of SA1 or SA2 in HeLa cells by siRNA has led to suggest that cohesin-SA2 is specifically required for centromeric cohesion whereas cohesin-SA1 is responsible for arm and telomere cohesion (Canudas and Smith, 2009).

*Corresponding author. Chromosome Dynamics Group, Molecular Oncology Programme, Spanish National Cancer Research Centre (CNIO), Melchor Fernández Almagro 3, 28029 Madrid, Spain.
Tel.: +34 917328000/ext. 3470; Fax: +34 917328033;
E-mail: alosada@cnio.es

Received: 12 September 2011; accepted: 9 January 2012; published online: 13 March 2012

To gain insight into the functional specificity of SA1 and SA2-containing cohesin complexes and their importance for cell viability, we have generated mouse embryos deficient for SA1. This approach ensures a complete absence of the protein and gives us the opportunity to transcend the cellular level in our studies and address the relevance of SA1 functions in embryonic development. This is particularly important in view of the existence of a human developmental disorder affecting 1 in 10 000 newborns, known as Cornelia de Lange syndrome (CdLS), caused by heterozygous mutations in the gene encoding the cohesin loader Nipbl/Scc2 (Liu and Krantz, 2009). Moreover, inactivation of SA2 has been very recently proposed to drive aneuploidy in human cancer (Solomon *et al*, 2011). Here, we show that SA1 deficiency results in delayed embryo development and lethality, which demonstrates that the functions of cohesin-SA1 and cohesin-SA2 are not redundant. We report for the first time that defects in cohesin-SA1 function generate aneuploidy and result in increased incidence and earlier onset of tumourigenesis. Unlike cohesin-SA2, lack of cohesin-SA1 leads to aneuploidy in the absence of precocious separation of sister chromatids (PSSC) in metaphase. Instead, chromosome segregation defects are caused by defective telomere replication in the absence of telomere cohesion, which is mediated exclusively by cohesin-SA1.

Results

Ablation of SA1 results in late embryonic lethality

We obtained from the German Gene Trap Consortium an embryonic stem (ES) cell line in which a rFlpROSA-geo cassette containing a splicing acceptor site and a polyadenylation sequence was inserted between exons 3 and 4 of the murine *Stag1* gene encoding the SA1-cohesin subunit (Figure 1A). The insertion creates a null allele and mouse embryos carrying this allele turn blue upon incubation in X-gal (Figure 1B). Heterozygous mice are viable and fertile but they do not produce homozygous progeny, indicating that SA1 function is essential for viability (Supplementary Table S1). Homozygous mutant embryos can be found at Mendelian ratios by day 11.5 of gestation (E11.5), allowing us to obtain mouse embryonic fibroblasts (MEFs) lacking SA1 mRNA (Figure 1C) or protein (Figure 1D). SA1 transcripts are 1–3 times more abundant than those of SA2 in MEFs, ES cells and in cells from different mouse tissues (Supplementary Figure S1A and B) whereas SA2 protein levels are three times higher than SA1 in MEFs (Supplementary Figure S1C). The levels of SA2 mRNA and SA2 protein remain virtually unchanged in the SA1-null MEFs (Figure 1C and D, respectively), suggesting that no compensatory upregulation of the other SA subunit occurs. SA1-null MEFs show clear proliferation defects (Figure 1E) that cannot be attributed to an obvious cell-cycle arrest or to senescence (Figure 1F and data not shown), and are highly aneuploid even in very early passages (Figure 1G). Importantly, in cells obtained from fetal livers at E14.5, without any passage in culture, aneuploidy is also higher in the SA1-null liver cells (Figure 1H). Viability of SA1-null embryos strongly decreases by E12.5 but some embryos that survive to E18.5 are much smaller than their wild-type littermates (Figure 1I) and show a substantial reduction of BrdU incorporation in most embryonic tissues (Figure 1J). They present a clear developmental delay and additional features that are reminiscent of CdLS, such as impaired lipid metabolism and delayed ossification (Remeseiro *et al*, 2012). Thus, some functions of cohesin-SA1 that cannot be performed by cohesin-SA2 are essential for cell proliferation and, thereby, to fulfil embryonic development.

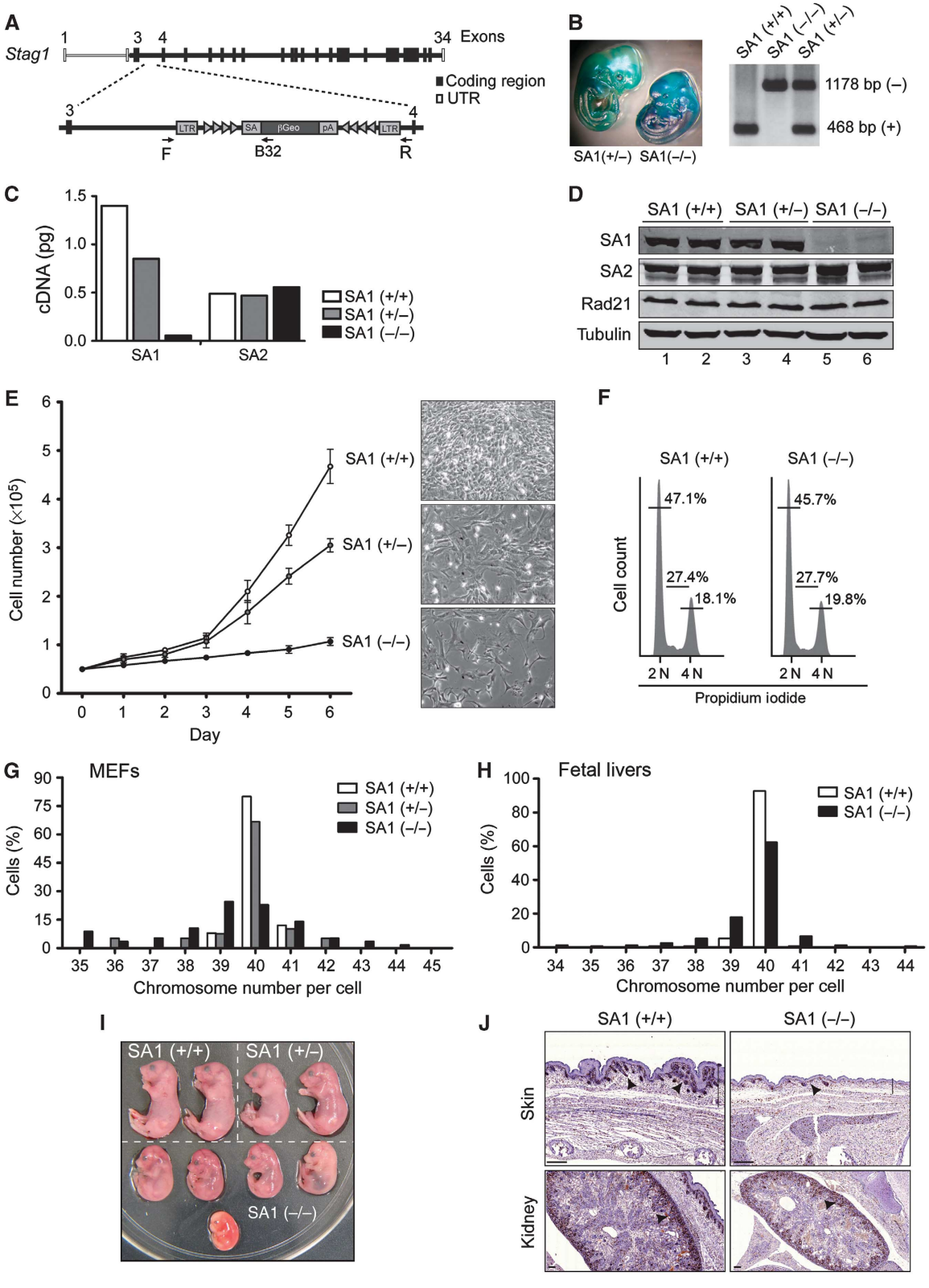
SA1 heterozygous animals have increased risk of cancer but are protected against acute tumourigenesis

SA1 heterozygous mice have increased incidence of spontaneous tumours, shorter lifespan than their wild-type littermates and some features of premature ageing such as acute kyphosis (Figure 2A and B; Supplementary Figure S2A, respectively). Monoallelic inactivation of SA1 accelerates the onset of tumourigenesis and changes the spectrum of tumours from that observed in wild-type mice. Increased incidence of hepatocellular carcinomas and some other carcinomas, as well as vascular tumours, is observed in SA1 heterozygous mice. Most remarkable is the early appearance of cystic papillary neoplasm (labelled as 'pancreas' in Figure 2A), which resembles human pancreatic intraductal papillary mucinous neoplasm and is extremely infrequent in mice (Supplementary Figure S2B, top panel). Tumours of mesenchymal origin, such as haematopoietic tumours (lymphomas and histiocytic sarcomas), are common in ageing wild-type mice of this genetic background, but appear earlier in the SA1 heterozygous animals (see examples in Supplementary Figure S2B). To further address the contribution of SA1 haploinsufficiency to tumourigenesis, we induced tumour formation both genetically and chemically. In terms of survival and tumour incidence, we did not observe significant differences between wild-type and SA1 heterozygous mice in p53-null and PTEN heterozygous backgrounds, probably due to the dominant effect that the loss of function of these tumour suppressors has in tumourigenesis (Supplementary Figure S2C). Strikingly, young SA1 heterozygous mice showed clear resistance to carcinogenesis upon treatment with 3-methyl-colanthrene (3-MC) and diethyl-nitrosamine (DEN) to induce fibrosarcomas and liver tumours, respectively (Figure 2C and E). Increased DNA damage signalling (γH2AX staining) and lower proliferation rates (Ki67 staining) could be observed in fibrosarcoma samples from SA1 heterozygous mice (Figure 2D; Supplementary Figure S2D). DEN-induced liver tumours from these animals also showed reduced number of Ki67-positive cells (Figure 2F). Thus, we suspect that the reduced proliferative capacity of SA1-deficient cells delays the growth of acute induced tumours in the young animals. However, SA1 haploinsufficiency increases susceptibility to spontaneous tumours and accelerates death in ageing mice by promoting aneuploidy.

Telomere-specific cohesion defects in SA1-null cells

Metaphase spreads from SA1-null cells present chromosomes with normally paired centromeres (Figure 3A), suggesting that cohesin-SA1 is not essential to maintain centromere cohesion in mouse chromosomes. Treatment of the mouse C2C12 cell line with siRNAs against SA1, SA2, both SA1 and SA2 or SMC1 further confirmed this result (Figure 3B; western blot analysis in Supplementary Figure S3A). The lack of only SA1 does not affect centromere cohesion, whereas the absence of SA2 does. Double depletion of SA1 and SA2, or depletion of SMC1 increases by almost four-fold the percen-

tage of metaphases with more than four chromosomes showing clear centromere cohesion defects (Figure 3B). This result suggests that both cohesin-SA1 and cohesin-SA2 contribute to arm cohesion so that in the absence of both separation of the sister centromeres becomes more evident. Since arm and telomere cohesion are difficult to assess in metaphase spreads, we performed instead fluorescence *in situ* hybridization (FISH) of interphase cells with probes from subtelomere and arm regions. Cohesion defects, revealed by the appearance of doublets, are significantly increased in the SA1-null MEFs for the subtelomeric probes, but not for the arm probes, suggesting that SA1 has a specific role in telomere cohesion in



mouse cells (Figure 3C). Downregulation of SA1 or SA2 in MEFs by siRNA also showed that telomere cohesion specifically depends on SA1 whereas arm cohesion in the loci examined was only affected by the absence of SA2 (Figure 3D). Taken all together, our results indicate that centromeric cohesion relies on cohesin-SA2 whereas cohesin-SA1 plays a unique role in telomere cohesion in mouse cells.

Aberrant telomeres in cells lacking SA1

We next looked at sister chromatid exchange at telomeres (T-SCE), an event that needs cohesion, by means of a FISH technique involving the use of telomeric probes specific for the lagging and the leading strands, chromosome orientation FISH or CO-FISH (Bailey *et al*, 2004). Consistent with the loss of cohesion, we found a two-fold decrease in T-SCE in cells lacking SA1 (Figure 4A). Telomere length was similar in wild-type and SA1-null MEFs (Figure 4B), whereas the number of telomere fusions was higher in the absence of SA1 (1.2 fusion events per 100 chromosomes compared with 0.17 in wild-type cells, $n=2$ clones per genotype, at least 1100 chromosomes counted). More striking was the high incidence of aberrant telomere structures in SA1-null mitotic chromosomes, as judged by the appearance of an irregularly shaped, not round but elongated or split telomeric FISH signal (Figure 4C and D). SA1 heterozygous MEFs showed an intermediate phenotype (Figure 4D, grey bar). These aberrant structures, known as fragile telomeres, have been described in mouse cells deficient for components of the shelterin complex that binds the TTAGGG repeats and safeguards the ends of mammalian chromosomes (Munoz *et al*, 2005; Palm and de Lange, 2008; Martinez *et al*, 2009; Sfeir *et al*, 2009). They receive this name for their resemblance to common fragile sites, chromosomal regions that challenge the replication machinery and are often visualized as breaks or gaps in metaphase chromosomes upon partial inhibition of DNA synthesis (Durkin and Glover, 2007). Indeed, treatment of wild-type MEFs with low doses of the DNA replication inhibitor aphidicolin increases significantly the frequency of fragile telomeres in control cells and even in SA1 heterozygous cells (Figure 4E, white and grey bars, respectively). However, this is not the case for SA1-null cells, maybe because a further increase in the already high incidence of fragile telomeres prevents progression to mitosis (Figure 4E, black bars). Telomere fragility in SA1-null cells does not result from impaired recruitment of shelterin to telomeres, as shown by both chromatin fractionation and ChIP-dot blot assays (Figure 4F and G). Importantly, expression of full-length SA1 in SA1-null MEFs rescues telomere fragility almost completely (Supplementary Figure S4), confirming that the telomere defect is due to impaired cohesin-SA1 function.

Deficient telomere replication in SA1-null cells

In order to assess telomere replication, we performed single-molecule analysis of replicated DNA (SMARD) on telomeric regions (Norio and Schildkraut, 2001; Sfeir *et al*, 2009) in primary wild-type and SA1-null MEFs. By counting the number of telomeric molecules containing a labelled nucleotide (IdU and/or CldU), we observed a highly significant reduction in the fraction of replicating telomeres in the SA1-null MEFs ($P<0.0001$; Figure 5A and B). Among the labelled telomeres, twice as many molecules were doubly labelled in the sample from SA1-null cells (13.9%) compared with wild-type cells (6.6%), which also reflects slower replication. Cell-cycle profiles of the two populations and BrdU incorporation at the time point studied were identical (Supplementary Figure S5), suggesting that there is not a general defect in S-phase progression. The specificity of this defect at telomeres is further supported by the SMARD analysis of the IgH locus as a control, which showed no evidence for altered replication dynamics in the absence of SA1 ($P=0.78$; Figure 5B and C). Thus, SA1 is specifically required for proper replication of telomeres.

Cohesion mediated by cohesin-SA1 contributes to efficient replication of telomeres

We next asked what is the mechanistic contribution of cohesin-SA1 to telomere replication. On one hand, cohesin affects origin activation most likely by promoting the proper organization of replication factories by intra-chromatid looping (Guillou *et al*, 2010). On the other, cohesion mediated by cohesin could contribute to HR required to restart stalled forks frequently found at telomeres (Sfeir *et al*, 2009; Badie *et al*, 2010). To distinguish between these two possibilities, we checked the effect of Sororin on telomere fragility. Sororin is a cohesin-interacting factor required for sister chromatid cohesion but not for origin activation (Guillou *et al*, 2010; Nishiyama *et al*, 2010). As in MEFs, telomere fragility occurred specifically upon treatment of C2C12 cells with siRNAs against SA1, but not SA2 (Figure 5D, grey bars), or upon addition of aphidicolin (Figure 5D, black bars). Importantly, cells with reduced Sororin showed a significant increase in telomere fragility (see Supplementary Figure S3B

Figure 1 A knockout mouse model for cohesin-SA1 subunit. (**A**) Schematic representation of the *Stag1*-knockout (KO) allele used in this study. The murine *Stag1* locus encoding SA1 contains 34 exons. The precise location of the gene trap cassette is indicated as well as the position of the primers used for genotyping. LTR, long terminal repeat; SA, splice acceptor; βgeo, β-galactosidase/neomycin phosphotransferase fusion gene; pA, polyadenylation sequence; triangles represent target sites for FLPe and Cre recombinases (Schnutgen *et al*, 2005). (**B**) X-gal staining of whole embryos carrying the KO allele in heterozygosis (left) and homozygosis (right), and PCR analysis of DNA purified from cells of the indicated genotypes. (**C**) Quantitative RT–PCR analysis to evaluate the mRNA levels of SA1 and SA2 in the indicated MEFs. Values are given as picograms per 2 μg of total RNA. (**D**) Western blot analysis of whole-cell extracts prepared from MEFs. Tubulin is used as loading control. (**E**) Growth curves of primary MEFs of the indicated genotypes (left) and representative images of the cultures by day 6 after plating the same number of cells. (**F**) DNA content profile of asynchronous wild-type and SA1-null primary MEFs. Percentage of cells in each phase of the cell cycle is shown. (**G**) Graph showing the distribution in the number of chromosomes of at least 100 metaphases from two clones of wild-type, two clones of heterozygous and six clones of SA1-null primary MEFs. (**H**) Same analysis carried out in cells from fetal livers. At least 150 metaphases from three wild-type and three SA1-null embryos were examined. (**I**) E18.5 embryos from the same litter were photographed and genotyped. Notice the reduced size of the SA1-null embryos. (**J**) BrdU staining of skin and kidney sections from E17.5 embryos of the indicated genotypes. Arrowheads point to proliferative areas, such as hair follicles in the skin (top panels) and the outer layer of the kidney (bottom panels). Automated quantification of the relative BrdU-positive area in whole embryo sections with Definiens Software shows a clear reduction of 23.4 ± 0.7% in SA1-null embryos with respect to wild-type ($n=4$ wild-type embryos and $n=5$ SA1-null E17.5 embryos were analysed). Scale bars, 200 μm (top) and 100 μm (bottom).

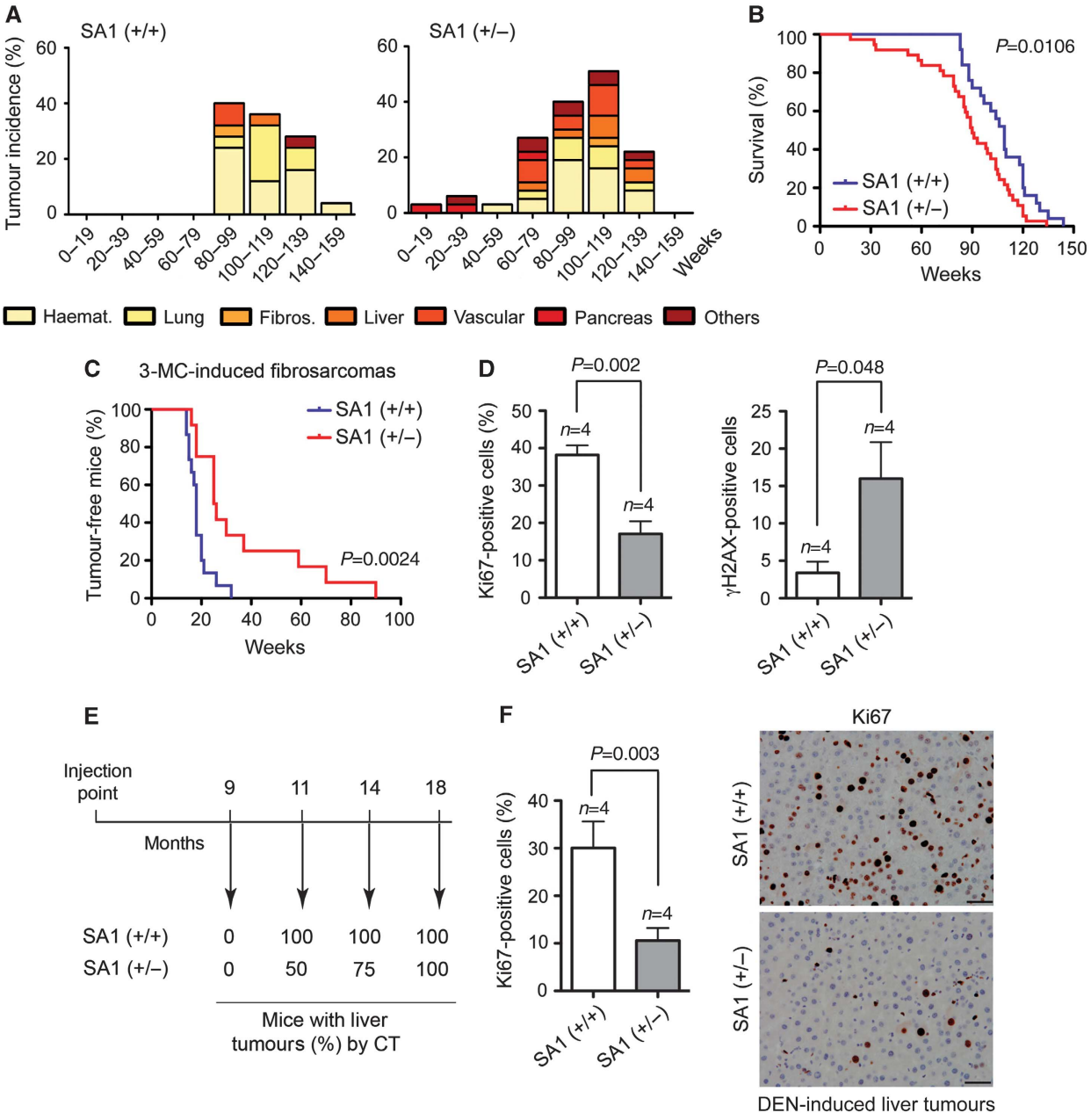


Figure 2 Reduced lifespan and increased incidence of spontaneous tumours in SA1 heterozygous mice, but higher resistance against chemically induced tumours. (**A**) Tumour incidence in wild-type ($n=25$) and SA1 heterozygous mice ($n=37$) relative to animal age in weeks. Note that SA1 heterozygous mice present higher tumour incidence and earlier onset of tumourigenesis. Haematopoietic tumours, fibrosarcomas, lung, liver, vascular and pancreas tumours are represented; the category 'others' includes osteomas, papillomas and mammary gland tumours. (**B**) Kaplan–Meier survival curves for wild-type (blue) and SA1 heterozygous mice (red) ($n=25$ and 37, respectively). (**C**) Kaplan–Meier curves showing tumour-free survival for wild-type and SA1 heterozygous mice injected with 3-MC to induce fibrosarcomas ($n=15$ animals per genotype). (**D**) Quantification of cells showing positive staining for Ki67 and γH2AX on tissue sections from fibrosarcomas like those shown in Supplementary Figure S2D. Five fields were counted per mouse of a total of four mice of each genotype. (**E**) Wild-type and SA1 heterozygous 15-day-old male mice ($n=4$ each) were injected with DEN and appearance of liver tumours was assessed by computed tomography (CT) at the indicated times after injection. (**F**) Quantification of Ki67-positive cells (left) in tissue sections of the liver tumours induced by DEN (right). Five fields were counted per mouse of a total of four mice of each genotype. Scale bars, 50 μm.

for western blot analysis). Thus, the cohesive function of cohesin-SA1 is required for efficient fork progression through subtelomere/telomere regions. In addition, we observed that depletion of SA1 or SA2 alone did not increase the percentage of chromosomes with breaks upon aphidicolin treatment compared with control cells, whereas simultaneous depletion of both or depletion of Sororin did (Figure 5E, black bars). This result indicates that both cohesin-SA1 and cohesin-SA2 contribute to cohesion and prevent fragile site breakage along chromosome arms, while only cohesin-SA1 can perform this function at telomeres.

Defective chromosome segregation in SA1-null cells leads to aneuploidy

To explore the mechanism that generates aneuploidy, we followed progression through mitosis of wild-type and

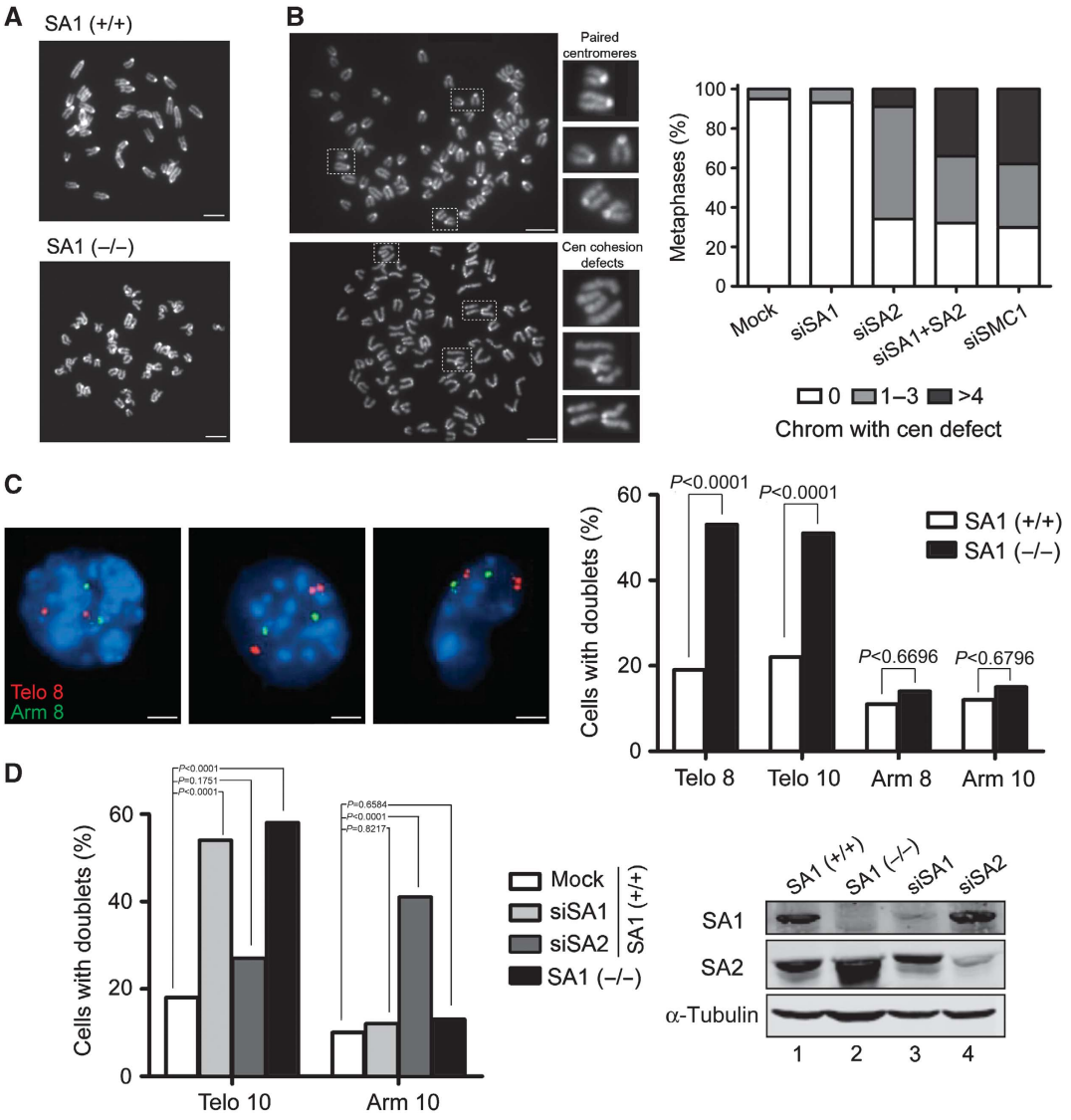
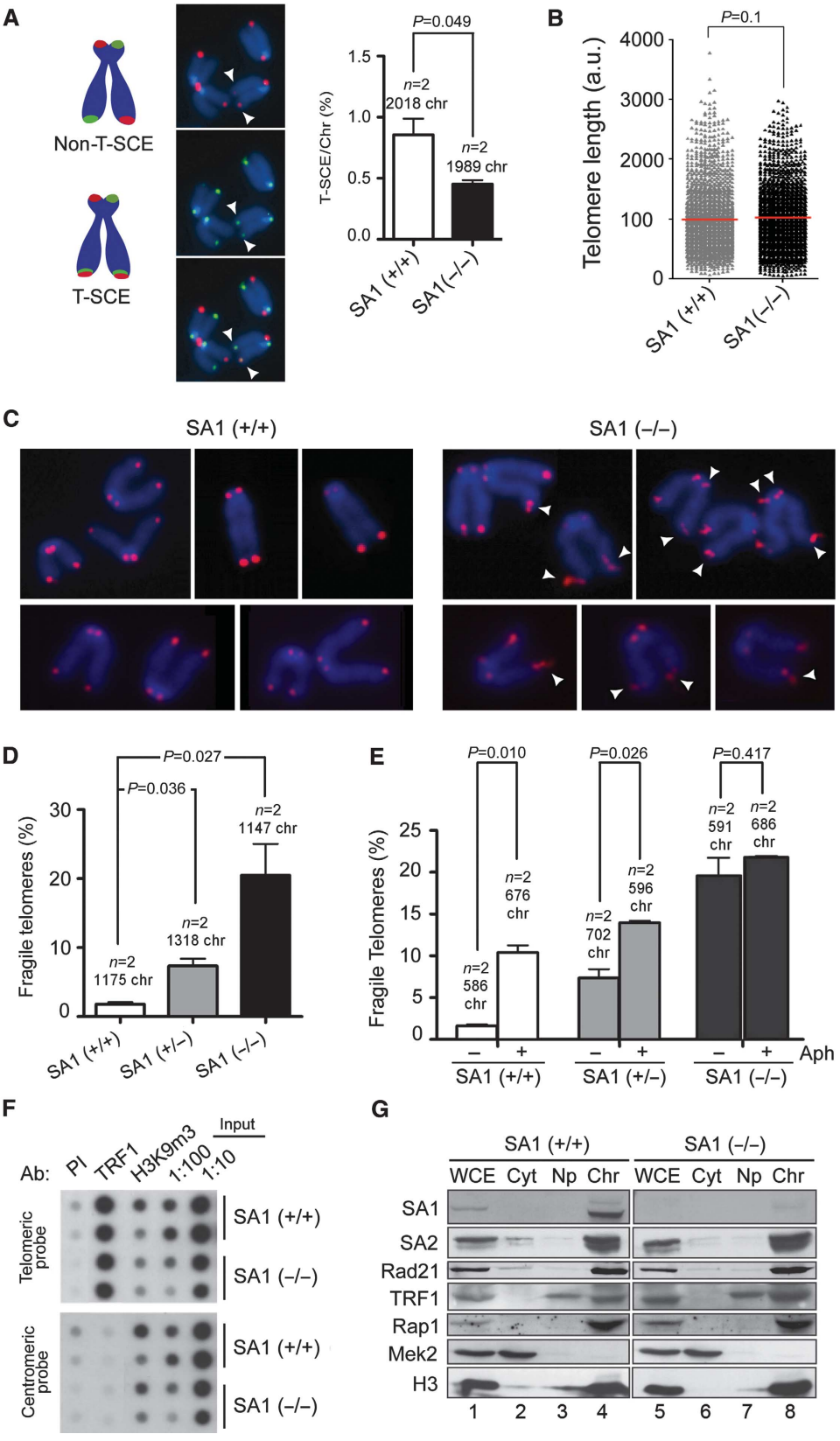


Figure 3 A specific role of SA1 in telomere cohesion. (**A**) Metaphase spreads from wild-type and SA1-null MEFs showing proper centromere cohesion. Scale bars, 10 μm. (**B**) Metaphase spreads from mouse C2C12 cells treated with control (top panel) and siRNA against SA2 (bottom panel) are shown as examples of proper and defective centromere cohesion, respectively. Scale bars, 10 μm. On the right, bar graph with the quantification of metaphase cells showing none, 1–3 or ⩾4 chromosomes with split centromeres after treatment with the indicated siRNAs ($n \geqslant 200$ metaphases per condition from two independent experiments). (**C**) FISH analysis of wild-type or SA1-null primary MEFs in interphase with probes from the subtelomeric regions of chromosome 8 and 10 (Telo 8 and Telo 10) and arm regions of the same chromosomes (Arm 8 and Arm 10). $n \geqslant 100$ cells per clone from two independent clones per genotype. Scale bars, 5 μm. (**D**) FISH analysis performed as in (**C**) in wild-type primary MEFs untransfected or transfected with siRNAs against SA1 (siSA1) or SA2 (siSA2). $n \geqslant 100$ cells per condition.

SA1-null MEFs by live-cell imaging. Previous results in human cells had shown that depletion of cohesin causes a Mad2-dependent metaphase arrest triggered by the lack of tension in the absence of proper centromeric cohesion (Toyoda and Yanagida, 2006). In contrast, the time that cells spend in mitosis before starting anaphase is similar in wild-type and SA1-null MEFs (median is 40 and 45 min, respectively; Figure 6A, blue lines), consistent with the finding of robust centromeric cohesion in these cells. However, almost 30% of mitotic cells lacking SA1 fail to complete mitosis appropriately and either are unable to complete cytokinesis and end up as binucleated cells (Figure 6A, green lines; Figure 6B, middle), or start anaphase but then collapse into one single tetraploid nucleus (Figure 6A, yellow lines), or die in mitosis before decondensing their chromosomes (mitotic catastrophe; Figure 6A, black lines; Figure 6B, bottom). Thus, cells in the two former categories become tetraploid after traversing mitosis. Staining of fixed cells also showed a five-fold increase in the number of binucleated cells among SA1-null cells compared with wild-type cells (Supplementary Figure S6). We have observed increased frequency of both lagging chromosomes and chromatin bridges among the SA1-null cells (Figure 6C, black bars). The presence of these structures could account for the observed cleavage furrow regression that results in either mitotic catastrophe or tetraploidization. While the former phenotype may contribute to the decreased proliferation rates of SA1-null cells, the latter is a reported initiator of aneuploidy (Ganem *et al*, 2007). We also found increased frequency of defective

anaphases in SA1 heterozygous cells, both *ex vivo* (in MEFs; Figure 6C, grey bars) and *in vivo* (in tissue sections; Supplementary Figure S7). This finding further supports the idea that aneuploidy resulting from faulty chromosome segregation promotes tumourigenesis in SA1 heterozygous animals.



Telomere replication defects cause chromosome missegregation in SA1-null cells

We next tested whether replicative stress could lead to chromosome segregation defects in MEFs. Indeed, wild-type MEFs treated with low doses of aphidicolin show high incidence of both lagging chromosomes and anaphases bridges (Figure 7A). Since SA1-deficient cells present replication defects specifically at telomeres, it is reasonable to propose that the chromosome segregation defects described in the previous paragraph result from impaired telomere replication. Immunostaining of SA1-null MEFs going through anaphase reveals that out of 21 anaphases displaying chromatin bridges none showed centromere (ACA) staining at the bridge, whereas telomere (TRF1) signals could be detected in 13 of them (Figure 7B, top and middle panels). The lack of TRF1 staining in the other eight anaphases could reflect loss of the shelterin upon chromatin stretching. This result supports the hypothesis that unresolved replication intermediates from subtelomere/telomere regions interfere with chromosome segregation and lead to the formation of anaphase bridges.

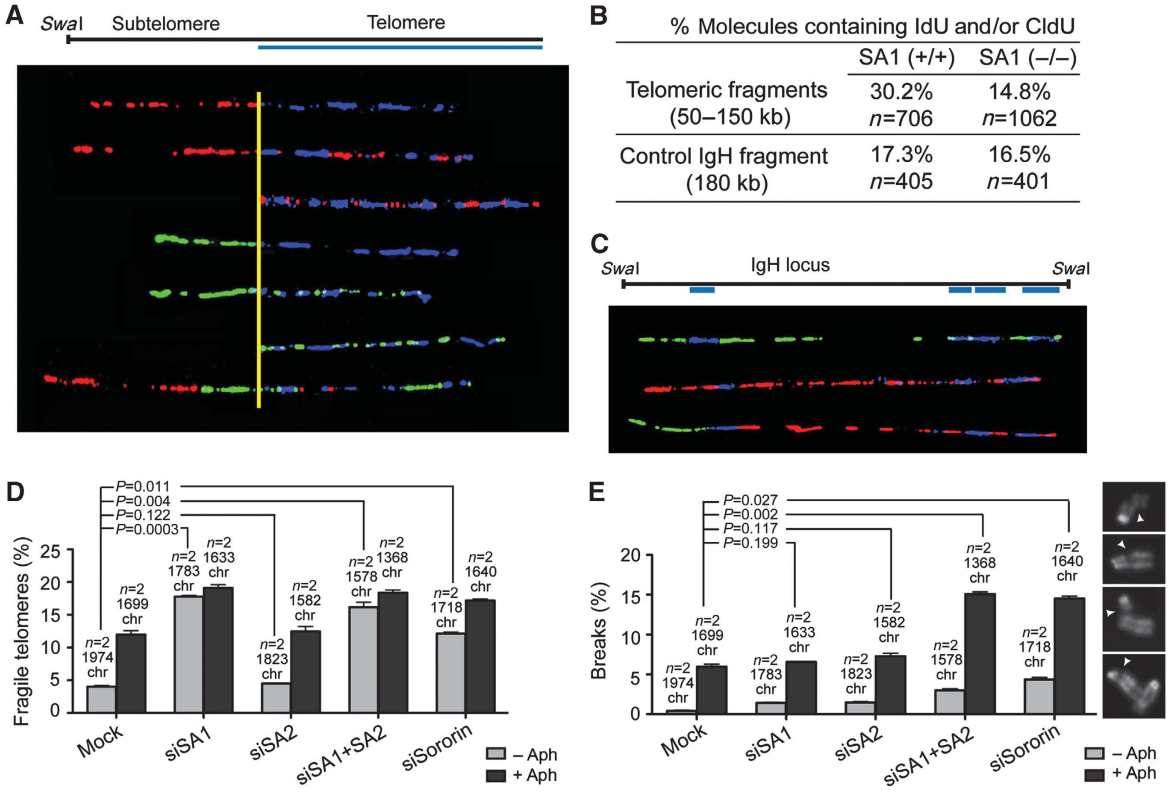


Figure 5 SMARD analysis of telomeric DNA reveals defective replication in SA1-null MEFs upon loss of telomere cohesion. (**A**) Examples of stretched telomeric DNA molecules of variable lengths (50–150 kb), identified by FISH with a telomeric probe (blue), that incorporated IdU (red) and/or CldU (green) during the time of the pulses. (**B**) Results of the SMARD analysis for telomeres (top) and molecules containing the IgH locus as control (bottom) from wild-type and SA1-null MEFs. (**C**) Examples of IdU/CldU (red/green) incorporation patterns in *Swa*I-digested fragments (180 kb) corresponding to the IgH locus, as identified by FISH with the indicated probes (blue). (**D**) Quantification of fragile telomeres in mouse C2C12 cells treated with no siRNA (mock) or siRNAs against SA1 (siSA1), SA2 (siSA2) or both (siSA1 + SA2) or Sororin (siSororin), either in the absence (grey bars) or in the presence of aphidicolin (black bars) as in Figure 4E. (**E**) Quantification of breaks along the chromosome arms in the indicated cells either untreated (grey bars) or treated with aphidicolin (black bars). The images on the right show examples of the broken chromosomes.

Figure 4 Telomere fragility in the absence of SA1. (**A**) T-SCE measured by CO-FISH (chromosome orientation FISH) in the telomeres of primary MEFs. The drawing on the left explains how T-SCE are visualized. The images on the middle show an example in which the telomeres indicated by arrowheads have undergone exchange since they are labelled by both the leading and the lagging strand-specific telomeric probes (green and red, respectively). Quantification of exchange events is shown in the bar graph on the right. (**B**) Telomere length measured by Q-FISH (quantitative FISH) for telomeres from wild-type and SA1-null MEFs (12 metaphase cells from two clones for each genotype). (**C**) Metaphase chromosomes from wild-type and SA1-null MEFs stained with a telomeric repeat probe (red) and DAPI (blue). Arrowheads point to fragile telomeres. (**D**) Quantification of fragile telomeres in chromosomes from two clones each of wild-type, SA1 (+/−) and SA1-null MEFs. (**E**) Telomere fragility measured in wild-type, SA1 (+/−) and SA1-null MEFs either untreated (−) or treated (+) with 0.5 μM aphidicolin for 24 h. In all cases two clones of each genotype were used. (**F**) ChIP-dot blot analysis of two clones each of wild-type and SA1-null MEFs with preimmune serum as negative control (PI), anti-TRF1 and anti-H3K9m3 as positive control. The chromatin obtained was transferred to a membrane and hybridized with a telomeric probe and a centromeric probe (major satellite) as control. TRF1 is present only at telomeres and its abundance is not affected by the lack of SA1. (**G**) Chromatin fractionation followed by immunoblotting with antibodies against cohesin subunits and shelterin proteins in wild-type and SA1-null cells. Mek2 cytoplasmatic kinase and histone H3 are used as control for the fractionation procedure. WCE, whole-cell extract; Cyt, cytoplasm; Np, nucleoplasm; Chr, chromatin fraction. The amount of TRF1 and Rap1 in chromatin does not depend on SA1 (lanes 4 and 8).

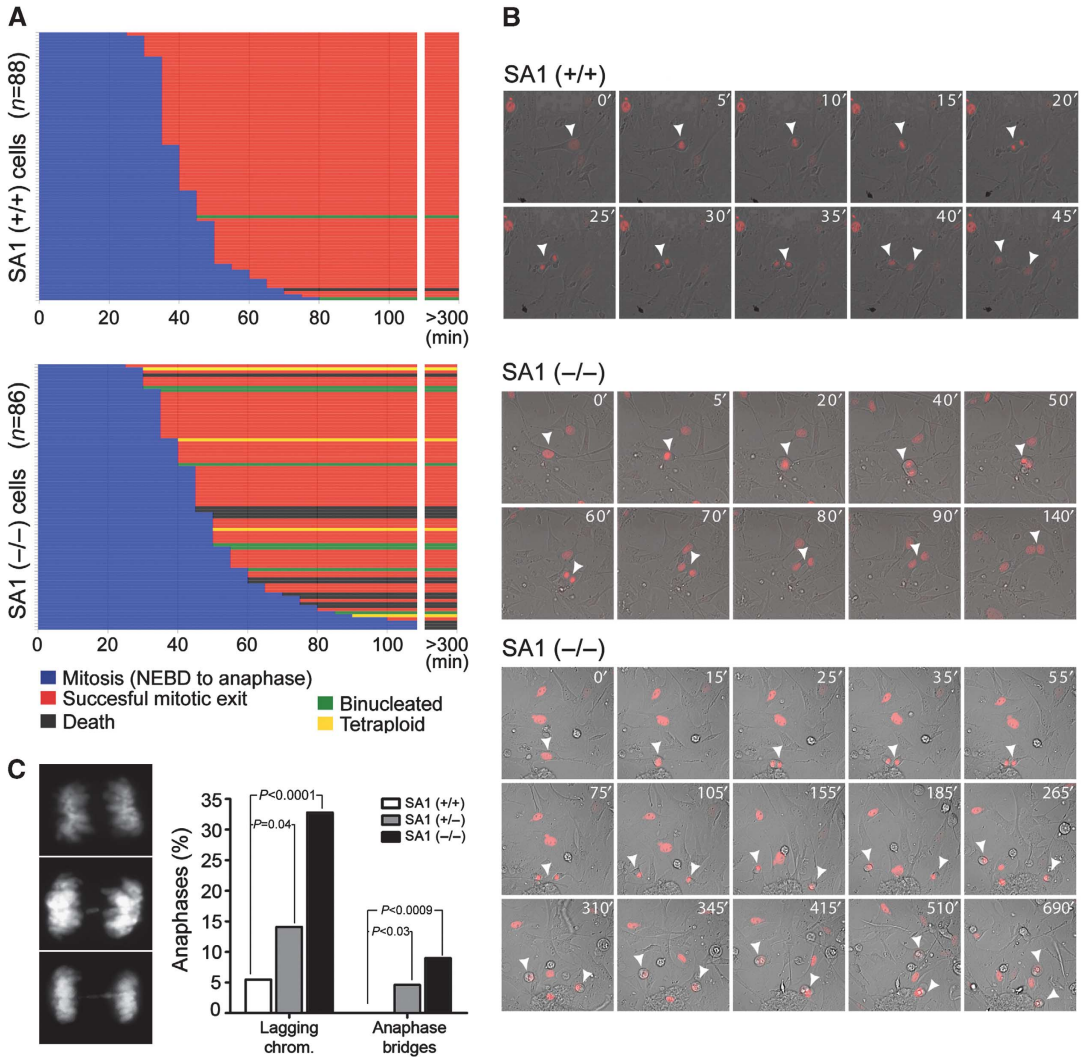


Figure 6 Chromosome segregation defects in SA1-null cells. (**A**) Graphical summary of the fates of mitotic cells from wild-type ($n=88$) and SA1-null MEFs ($n=86$) observed by live-cell imaging. Each line represents the progression through mitosis of a single cell and it is coloured according to the legend shown below the graph. (**B**) Examples of wild-type and SA1-null MEFs progressing through mitosis after transfection with H2B-mCherry (red) to label chromatin. The time after NEBD for each frame is indicated. Examples of a normal mitosis (top), a mitosis that leads to the formation of a binucleated cell (middle) and a mitosis that ends up in cell death (bottom) are shown. (**C**) Aberrant anaphases in SA1 (+/−) and SA1-null MEFs. Examples of a proper anaphase (top), an anaphase with a lagging chromosome (middle) and an anaphase with a chromatin bridge (bottom) are shown ($n\geqslant 50$ cells per clone from two independent clones per genotype).

We also examined anaphases with lagging chromosomes and found that in all cases (20 out of 20) the lagging chromosome consisted of two paired chromatids (Figure 7B, bottom panels). In this case, we speculate that unreplicated regions in the short arm telomere, which is next to the centromere, prevent the separation of sister centromeres.

Taken all together, we show that the absence of telomere cohesion in cells lacking SA1 results in defective telomere replication. This generates fragile telomeres and is the cause of chromosome missegregation leading to mitotic catastrophe and aneuploidy.

Discussion

A mouse model to address cohesin functions in vivo

The importance of cohesin function in cell proliferation and differentiation is expanding rapidly (Merkenschlager, 2010; Dorsett, 2011). Mouse models appear as appropriate tools to unravel these different functions and their regulation both temporally (during embryo development and adult ageing) and spatially (in different tissues and organs). To date, mice deficient for genes encoding components of the cohesin complex have been generated for three meiosis-specific subunits, Smc1 β (Revenkova *et al*, 2004), Rec8 (Xu *et al*, 2005) and Rad21L (Herran *et al*, 2011) and also recently for Rad21 (Xu *et al*, 2010). The first three are viable but sterile whereas biallelic inactivation of Rad21 results in early embryonic lethality and only the radiosensitivity of Rad21 heterozygous animals was reported. A mouse with a conditional allele of the Rad21 locus has also been generated and used for depleting cohesin specifically in non-cycling thymocytes (Seitan *et al*, 2011). Thus, our study presents the first thorough characterization of a mouse model of a ubiquitously expressed cohesin subunit, SA1, for embryonic development and cancer. SA1-null embryos are embryonic lethal, which demonstrates that cohesin-SA1 performs a function that

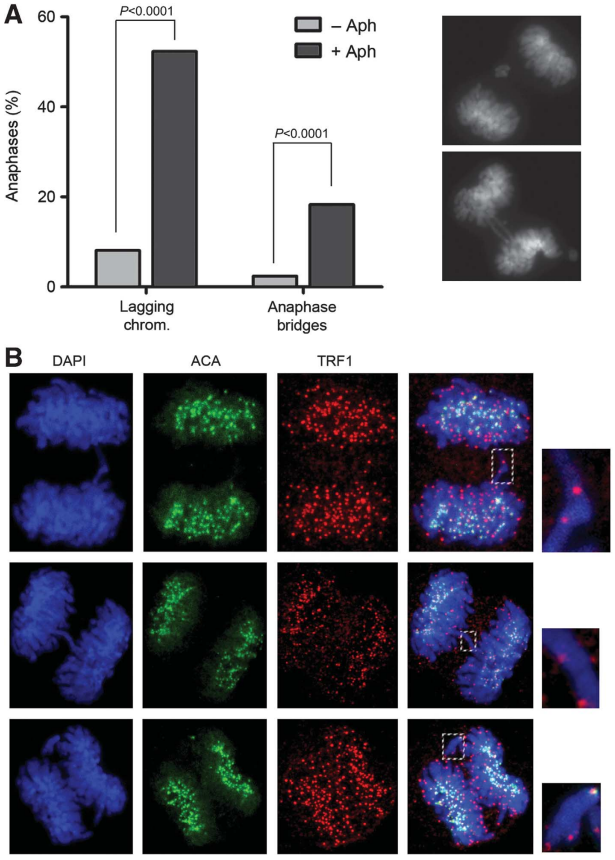


Figure 7 Defective telomere replication causes chromosome mis-segregation. (**A**) Frequency of anaphase bridges and lagging chromosomes in wild-type MEFs untreated (–Aph) or treated (+Aph) with low doses of aphidicolin (0.5 μM 24 h). Examples of chromosome segregation defects induced by global inhibition of replication are shown on the right. (**B**) Examples of SA1-null anaphase cells with chromosome segregation problems stained with antibodies against TRF1 (red) and ACA (green). Cells were pre-extracted with detergent before fixation. The top and middle panels show the presence of telomeres but no centromeres at the chromatin bridges. The bottom panel shows a lagging chromosome containing two sister chromatids. Confocal microscopy was used to ensure that TRF1 signals were in the same focal plane as the DNA.

cannot be accomplished by cohesin-SA2. Inactivation of SA1 permits embryos to survive at least to E12.5 and, in some cases, even to E18.5. Decreased proliferation rates of the SA1-null cells due to metabolic stress resulting from aneuploidy (Williams *et al*, 2008), to reduced cell survival upon passage through mitosis (this study) or to transcriptional dysregulation of pro-proliferation genes like myc (Rubio *et al*, 2008; Rhodes *et al*, 2010; Remeseiro *et al*, 2012) likely contribute to the lethality of SA1-null embryos.

Cohesin-SA1 is required for efficient telomere replication

Cohesin complexes carrying SA1 and SA2 subunits are not functionally redundant. We have shown that cohesin-SA2 is more critical for centromeric cohesion and cohesin-SA1 has a specific role in telomere cohesion, consistent with previous results in HeLa cells (Canudas and Smith, 2009). Whereas both contribute to cohesion and prevent fragile site formation along chromosome arms, only cohesin-SA1 can perform this role at telomeres. We show mechanistic evidence of the deleterious effects of cohesion loss for efficient replication of telomeres, since downregulation of Sororin, a cohesin-interacting factor specifically required for cohesin to become cohesive, results also in telomere fragility. The repeated nature of telomeric sequences, the propensity of their G-rich strand to form G4 DNA structures, and the presence of the t-loop hinder the passage of replication forks (Gilson and Geli, 2007). Furthermore, stalled forks occurring at chromosome ends cannot be rescued by forks progressing in the opposite direction. Recent reports shed light on how the cell deals with this problem (Sfeir *et al*, 2009; Badie *et al*, 2010; Ye *et al*, 2010). Replicative stress specifically at telomeres increases dramatically in the absence of SA1, as evidenced by the high incidence of fragile telomeres and confirmed by the results of the single-molecule analyses (SMARD). We propose that cohesion mediated by cohesin-SA1 is an important component of the pathway that stabilizes arrested forks at telomeres and facilitates their restart and/or promotes HR-mediated repair of breaks generated upon fork collapse.

A new mechanism to generate aneuploidy when cohesin function is impaired

We have shown that SA1-null cells, despite their robust centromere cohesion, have defects in chromosome segregation that most likely arise from defective telomere replication (see model in Figure 8). Aberrant telomere structures resulting from incomplete replication might hinder proper orientation of the sister kinetochores in the acrocentric mouse chromosomes, thus increasing the chance of merotely (i.e., a single kinetochore attached to both spindle poles), which generates lagging chromatids in anaphase (Cimini *et al*, 2001; Salmon *et al*, 2005). A close analysis of defective anaphases in SA1-null cells shows that lagging chromosomes consist of two paired sister chromatids, not just one. We therefore favour the alternative possibility that unreplicated regions in subtelomere/telomere regions of the short arm prevent the separation of sister centromeres. If the unresolved replication intermediate happened in the long arm, it would generate instead a chromatin bridge containing telomere sequences, and we show evidence for these as well. Our hypothesis is consistent with recent reports showing that incompletely replicated regions or unresolved replication intermediates, or disruption of homology-directed repair, give rise to chromatin bridges and lagging chromosomes in anaphase (Torres-Rosell *et al*, 2007; Chan *et al*, 2009; Naim and Rosselli, 2009; Kawabata *et al*, 2011; Laulier *et al*, 2011; Lukas *et al*, 2011). The presence of these structures prevents proper cytokinesis and results in either mitotic catastrophe or tetraploidization, and the latter could drive aneuploidy in SA1-deficient cells. Ablation of telomere proteins can also lead to aneuploidy by additional mechanisms involving either chromosome fusion–breakage–bridge cycles following telomere de-protection (Smogorzewska *et al*, 2002) or bypass of mitosis upon persistent damage signalling in the uncapped telomeres (Davoli *et al*, 2010).

Importantly, SA1 heterozygosity also leads to telomere fragility, chromosome segregation defects and aneuploidy. This aneuploidy is likely to contribute to tumourigenesis, as we show that SA1 heterozygous mice develop spontaneous tumours earlier than their wild-type littermates and with a distinct tumour spectrum. Until recently, little was known of the involvement of cohesin mutations in human cancer (Xu

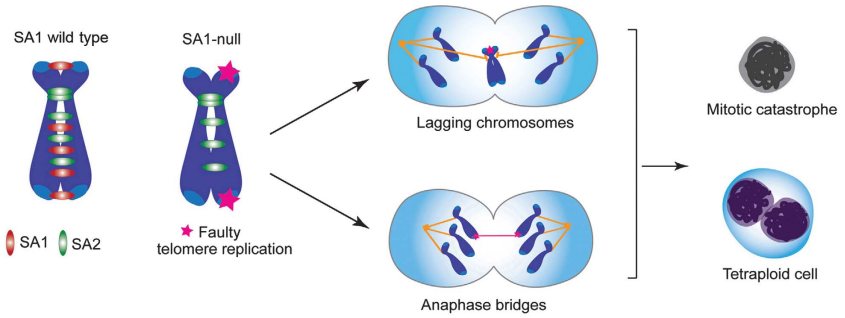


Figure 8 The role of cohesin-SA1 in telomere cohesion and replication is essential for accurate chromosome segregation and to prevent aneuploidy. See text for details on the model.

et al, 2011). Inactivation of SA2 has now been found in a diverse range of tumour types and it has been proposed that lack of SA2 function leads to aneuploidy due to PSSC in metaphase (Solomon *et al*, 2011), in agreement with previous reports (McGuinness *et al*, 2005; Canudas and Smith, 2009). We propose that inactivation of SA1 also generates aneuploidy, but through a completely different mechanism involving impaired telomere replication.

Materials and methods

Generation of SA1-knockout mouse model, MEFs isolation and cell culture

An ES cell line containing a SA1-Gene Trap allele (PO99A04) was obtained from the German Gene Trap Consortium (Schnutgen *et al*, 2005). Chimeric mice were generated by microinjection of ES clones into S129V host blastocysts, which were then implanted into pseudopregnant C57Bl/6J females, and germline transmission was assessed. Mice were housed in a pathogen-free animal facility following the animal care standards of the institution. The following primers were used for genotyping by PCR: F, 5′-GTGCTAGGATG ACTCTGAAACTG-3; B32, 5′-CAAGGCGATTAAGTTGGGTAA CG-3′; R 5′-TGTGCTAGGCAGACAGTCTCC-3. Primary MEFs were isolated from E12.5 embryos resulting from intercrosses of SA1 heterozygous mice and cultured in DMEM/10% FBS.

Q-FISH, CO-FISH and interphase FISH

Metaphases for Q-FISH were prepared and hybridized as previously described (Martinez *et al*, 2009). Images were captured using Leica Q-FISH software and quantification of telomere intensities to determine telomere length was performed with TFL-Telo software. In the indicated cases, cells were cultured in 0.5 μM aphidicolin for 24 h. For CO-FISH, MEFs were cultured in the presence of 10 μM BrdU for 24 h to ensure complete replication and then CO-FISH was performed using first a fluorescein-labelled $(TTAGGG)_7$ probe and then a Cy3-labelled $(CCCTAA)_7$ probe (Applied Biosystems). FISH on interphase cells was done according to standard protocols, using probes from subtelomere and arm regions, which were labelled by nick-translation (Abbott Inc.) on BAC clones from a mouse library [BAC references: RP23-326G18 (telomere8), RP23-310L10 (arm8), RP23-71E10 (telomere10), RP23-453P21 (arm10)]. The percentage of cells with doublets was determined upon counting at least 100 cells per genotype and per experiment.

SMARD

For SMARD assay, asynchronously growing MEFs were sequentially labelled with 25 mM IdU (4 h) and 25 mM CldU (4 h). Cells were embedded in agarose plugs, lysed and DNA was digested with *Swa*I and fractionated by PFGE to select 50–150 kb fragments enriched in telomeric DNA as described (Norio and Schildkraut, 2001; Sfeir *et al*, 2009). DNA was stretched on microscope slides coated with 3-amino-propyltriethoxysilane (Sigma), denatured in alkali-denaturing buffer (0.1 N NaOH in 70% ethanol and 0.1% β-mercaptoethanol) for 12 min and fixed by addition of 0.5% glutaraldehyde for 5 min. Telomeric DNA was detected by hybridization with a Biotin-OO-$(CCCTAA)_4$ PNA probe and Alexa Fluor 350-conjugated NeutrAvidin antibody (Molecular Probes) followed by biotinylated anti-avidin antibody (Vector). Halogenated nucleotides were detected with a mouse anti-IdU monoclonal antibody (BD), a rat anti-CldU monoclonal antibody (Accurate) and Alexa Fluor-conjugated secondary antibodies.

Chromosome spreads from MEFs and fetal livers

MEFs in culture were arrested in 0.1 μg/ml colcemide for 4–6 h, harvested by trypsinization, swollen in 75 mM KCl for 30 min at 37°C and fixed in methanol:acetic acid 3:1. Fetal livers from E14.5 embryos were minced with a scalpel and the cell suspension was incubated for 10 min at 37°C in 0.1 μg/ml colcemide in EDTA-containing buffer, further incubated in 75 mM KCl for 15 min and fixed in methanol:acetic acid 3:1. In both cases, the fixed suspension was dropped onto slides to obtain chromosome spreads that were stained and mounted with ProLong-Gold with DAPI (Invitrogen) and visualized using a Leica fluorescence microscope.

FACS

FACS analysis for DNA content was performed using propidium iodide staining, according to the standard procedures. For BrdU staining, cells were pulsed with 10 μM BrdU for 40 min, fixed and incubated with FITC-conjugated anti-BrdU antibody (BD Biosciences). Flow cytometry was performed using the FACS Canto II (Becton Dickinson) and data were analysed using FlowJo software (version 9.3.1).

Quantitative RT–PCR

The amount of SA1 and SA2 transcripts in MEFs, ES cells and different mouse tissues was determined by absolute quantitative RT–PCR. Total RNA was extracted using the RNeasy Mini Kit (Qiagen) and retrotranscribed with Superscript II (Invitrogen) using random hexamer primers. An Applied Biosystems 7900HT Fast qRT–PCR was used to determine the mRNA levels. The following primers were used: SA1 (F: 5′AGGCTTTCATGCTGCTCTGT3′ and R: 5′TCCATGCTTT GGTTTTCCTC3′), SA2 (F: 5′GGGGGAGGAACTGT CTTTCT3′ and R: 5′CCTTCAATGTCTTCAAAATCTGTG3′) and GAPDH as reference gene (F: 5′TGCACCACCAACTGCTTAGC3′ and R: 5′GAGG GGCCATCCACAGTCTTC3′).

Extract preparation, immunoblotting, ChIP-dot blot and immunofluorescence

For whole-cell extracts, cells were collected by trypsinization, washed once in cold PBS, resuspended in SDS–PAGE loading buffer and sonicated. For protein extracts from organs, a piece of tissue was pulverized in a mortar containing liquid nitrogen and lysed in RIPA buffer. Equal amounts of protein were run in either 7.5 or 12.5% Bis/Tris gels followed by western blotting. Chromatin fractionation was performed according to the protocol described by Mendez and Stillman (2000). For ChIP-dot blot, immunoprecipitates were dot-blotted into Hybond N+ membrane and hybridization was performed with a radiolabelled telomeric probe recognizing the TTAGGG repeats, and a centromeric probe annealing with the mouse major satellite (Benetti *et al*, 2007). For immunofluorescence, cells were cultured on polylysine-coated coverslips, fixed in 4% formaldehyde for 15 min at room temperature, permeabilized in 0.2% Triton-X100 for 5 min and

subjected to antibody incubation. Images were taken using a Leica DM6000 microscope or with a confocal SP5-WLL (for Figure 7). Rabbit polyclonal sera against SA1, SA2, SMC1 and Sororin were obtained by using as immunogen either a synthetic peptide [SA1-C (CEDDSGFGMPMF), SA2 (CDPASIMDESVLGVSMF), SMC1 (CDLTKYPDANPNPNEQ)] or full-length recombinant protein (mSororin, transferred from IMAGE clone 30065848 into pDEST17) and affinity purified. Other custom made and commercial antibodies used in this study were as follows: Rad21 (Losada *et al*, 1998); TRF1 and Rap1 (Martinez *et al*, 2009, 2010); H3K9m3 (Millipore, 07-442); α-tubulin (Sigma, DM1A); GAPDH (Sigma, G8795); Histone 3 (Abcam, AB1791); Mek2 (BD, M24520); phalloidin-488 (Invitrogen); ACA (Antibodies Inc., 15-235).

RNA interference and rescue experiments

Interference of SA1, SA2, Sororin and SMC1 was performed with siGENOME SMARTpool siRNAs from Dharmacon (M-041989, M-057033, M-048366 and M-049483, respectively) at a final concentration of 100 nM and experiments were done 72–96 h after transfection. For interference of C2C12 cells and MEFs, DharmaFECT transfection reagent 1 (Dharmacon) and the Neon transfection system (Invitrogen) were used, respectively. Rescue experiments were performed by electroporating 10^6 cells with 12 μg of full-length SA1 cloned in pBABE-puro vector and telomere fragility was assessed 40 h later.

Recombinant protein expression

RNA obtained from MEFs was used for retrotranscription and subsequent PCR amplification of cDNA encoding SA1 and SA2 C-terminal regions (amino acids 1089–1258 and 992–1162, respectively). The resultant DNA fragments were then cloned into the *Eco*RI site of the pGEX-KG expression vector. Recombinant fusion proteins were then expressed in BL21-pLys strain and purified following standard methods. These proteins were used to quantitate the amount of SA1 and SA2 in MEFs (Supplementary Figure S1C).

Live-cell imaging

MEFs were transfected with H2B-mCherry expression vector using the Amaxa Nucleofector System and seeded onto chamber-slides. Time-lapse live-cell imaging was performed using the Delta Vision system (Applied Precision). Phase-contrast and fluorescent images were acquired every 5 min with a ×10 objective for 24 h. Image analysis was performed using ImageJ software.

Histology and immunohistochemistry

Embryos, normal tissues and tumour samples were fixed in 10% buffered formalin (Sigma) and embedded in paraffin using standard procedures. For histopathological studies, 3 μm sections were stained with haematoxylin and eosin (HE). Anti-BrdU (GE Healthcare), Ki67 (Master Diagnostica) and γH2AX (Millipore, 05-636) primary antibodies were used for immunohistochemical analysis, positive cells were visualized using 3,3-diaminobenzidine tetrahydrochloride plus (DAB+) as a chromogen and counterstaining was performed with haematoxylin.

Carcinogen treatments

For the induction of fibrosarcomas, 8-week-old mice received a single intramuscular injection of 1 mg of 3-MC diluted in 100 μl of sesame oil in one of the rear legs. Mice were observed on a daily basis until tumours of 1.5 cm in diameter developed in the injected leg, at which point the animals were sacrificed and the tumours were extracted for further analysis. Liver tumours were induced in 15-day-old male mice by intraperitoneal injection of 50 μg of DEN diluted in 100 μl of PBS. Tumour development was followed by computed tomography (CT). Animals were observed daily and sacrificed when they manifested signs of morbidity, in accordance with the Guidelines for Humane End Points for Animals used in biomedical research.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 5 software. A two-tailed Fisher's exact test was done in Figures 3C and D, 5B, 6C and 7A, and for the rest of statistical analysis a two-tailed Student's *t*-test was applied. Data are shown as mean ± s.e.m (standard error of the mean); $P<0.05$ were considered significant and actual *P*-values are depicted in the figures.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We are grateful to I Barthelemy for her initial work on this project, M Rodríguez-Corsino for excellent technical assistance and B Ferreira, MC Martín and JC Cigudosa (Cytogenetics Unit, CNIO) for providing the BACs and invaluable help with FISH. We also acknowledge M Malumbres, A Martín-Pendás and M Soengas for helpful discussions and O Fernández-Capetillo and M Serrano for critically reading the manuscript. This research has been supported by the Spanish Ministry of Science and Innovation (SAF-2010-21517 and CSD2007-00015 Inesgen/FEDER to AL; 'Ramón y Cajal' grants for AC and PM), a La Caixa predoctoral fellowship for SR and NIH Grant GM045751 to CLS.

Author contributions: AL designed and supervised the study. SR designed, performed, analysed and interpreted most of the experiments with contributions from AC in Figures 1, 3 and 5; Supplementary Figures S1 and S4; from MC^1 in Figure 1; and from PM and MB in Figure 4. SMARD was performed by SR in the laboratory of CLS with advise from WCD. MC^4 performed the histopathological analyses. AL and SR wrote the manuscript with ideas and comments from the other authors.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Anderson DE, Losada A, Erickson HP, Hirano T (2002) Condensin and cohesin display different arm conformations with characteristic hinge angles. *J Cell Biol* **156:** 419–424

Badie S, Escandell JM, Bouwman P, Carlos AR, Thanasoula M, Gallardo MM, Suram A, Jaco I, Benitez J, Herbig U, Blasco MA, Jonkers J, Tarsounas M (2010) BRCA2 acts as a RAD51 loader to facilitate telomere replication and capping. *Nat Struct Mol Biol* **17:** 1461–1469

Bailey SM, Brenneman MA, Goodwin EH (2004) Frequent recombination in telomeric DNA may extend the proliferative life of telomerase-negative cells. *Nucleic Acids Res* **32:** 3743–3751

Benetti R, Gonzalo S, Jaco I, Schotta G, Klatt P, Jenuwein T, Blasco MA (2007) Suv4-20 h deficiency results in telomere elongation and derepression of telomere recombination. *J Cell Biol* **178:** 925–936

Canudas S, Smith S (2009) Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells. *J Cell Biol* **187:** 165–173

Carramolino L, Lee BC, Zaballos A, Peled A, Barthelemy I, Shav-Tal Y, Prieto I, Carmi P, Gothelf Y, Gonzalez de Buitrago G, Aracil M, Marquez G, Barbero JL, Zipori D (1997) SA-1, a nuclear protein encoded by one member of a novel gene family: molecular cloning and detection in hemopoietic organs. *Gene* **195:** 151–159

Chan KL, Palmai-Pallag T, Ying S, Hickson ID (2009) Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat Cell Biol* **11:** 753–760

Cimini D, Howell B, Maddox P, Khodjakov A, Degrassi F, Salmon ED (2001) Merotelic kinetochore orientation is a major mechanism of aneuploidy in mitotic mammalian tissue cells. *J Cell Biol* **153:** 517–527

Davoli T, Denchi EL, de Lange T (2010) Persistent telomere damage induces bypass of mitosis and tetraploidy. *Cell* **141:** 81–93

Dorsett D (2011) Cohesin: genomic insights into controlling gene transcription and development. *Curr Opin Genet Dev* **21:** 199–206

Durkin SG, Glover TW (2007) Chromosome fragile sites. *Annu Rev Genet* **41:** 169–192

Gandhi R, Gillespie PJ, Hirano T (2006) Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Curr Biol* **16:** 2406–2417

Ganem NJ, Storchova Z, Pellman D (2007) Tetraploidy, aneuploidy and cancer. *Curr Opin Genet Dev* **17:** 157–162

Gause M, Misulovin Z, Bilyeu A, Dorsett D (2010) Dosage-sensitive regulation of cohesin chromosome binding and dynamics by Nipped-B, Pds5, and Wapl. *Mol Cell Biol* **30:** 4940–4951

Gilson E, Geli V (2007) How telomeres are replicated. *Nat Rev Mol Cell Biol* **8:** 825–838

Guacci V, Koshland D, Strunnikov A (1997) A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. *cerevisiae*. *Cell* **91:** 47–57

Guillou E, Ibarra A, Coulon V, Casado-Vela J, Rico D, Casal I, Schwob E, Losada A, Mendez J (2010) Cohesin organizes chromatin loops at DNA replication factories. *Genes Dev* **24:** 2812–2822

Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M (2009) Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* **460:** 410–413

Haering CH, Farcas AM, Arumugam P, Metson J, Nasmyth K (2008) The cohesin ring concatenates sister DNA molecules. *Nature* **454:** 297–301

Hauf S, Roitinger E, Koch B, Dittrich CM, Mechtler K, Peters JM (2005) Dissociation of cohesin from chromosome arms and loss of arm cohesion during early mitosis depends on phosphorylation of SA2. *PLoS Biol* **3:** e69

Herran Y, Gutierrez-Caballero C, Sanchez-Martin M, Hernandez T, Viera A, Barbero JL, de Alava E, de Rooij DG, Suja JA, Llano E, Pendas AM (2011) The cohesin subunit RAD21L functions in meiotic synapsis and exhibits sexual dimorphism in fertility. *EMBO J* **30:** 3091–3105

Holzmann J, Fuchs J, Pichler P, Peters JM, Mechtler K (2010) Lesson from the stoichiometry determination of the cohesin complex: a short protease mediated elution increases the recovery from cross-linked antibody-conjugated beads. *J Proteome Res* **10:** 780–789

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467:** 430–435

Kawabata T, Luebben SW, Yamaguchi S, Ilves I, Matise I, Buske T, Botchan MR, Shima N (2011) Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. *Mol Cell* **41:** 543–553

Kueng S, Hegemann B, Peters BH, Lipp JJ, Schleiffer A, Mechtler K, Peters JM (2006) Wapl controls the dynamic association of cohesin with chromatin. *Cell* **127:** 955–967

Lafont AL, Song J, Rankin S (2010) Sororin cooperates with the acetyltransferase Eco2 to ensure DNA replication-dependent sister chromatid cohesion. *Proc Natl Acad Sci USA* **107:** 20364–20369

Laulier C, Cheng A, Stark JM (2011) The relative efficiency of homology-directed repair has distinct effects on proper anaphase chromosome separation. *Nucleic Acids Res* **39:** 5935–5944

Liu J, Krantz ID (2009) Cornelia de Lange syndrome, cohesin, and beyond. *Clin Genet* **76:** 303–314

Losada A, Hirano M, Hirano T (1998) Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. *Genes Dev* **12:** 1986–1997

Losada A, Yokochi T, Hirano T (2005) Functional contribution of Pds5 to cohesin-mediated cohesion in human cells and Xenopus egg extracts. *J Cell Sci* **118:** 2133–2141

Losada A, Yokochi T, Kobayashi R, Hirano T (2000) Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. *J Cell Biol* **150:** 405–416

Lukas C, Savic V, Bekker-Jensen S, Doil C, Neumann B, Pedersen RS, Grofte M, Chan KL, Hickson ID, Bartek J, Lukas J (2011) 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat Cell Biol* **13:** 243–253

Martinez P, Thanasoula M, Carlos AR, Gomez-Lopez G, Tejera AM, Schoeftner S, Dominguez O, Pisano DG, Tarsounas M, Blasco MA (2010) Mammalian Rap1 controls telomere function and gene expression through binding to telomeric and extratelomeric sites. *Nat Cell Biol* **12:** 768–780

Martinez P, Thanasoula M, Munoz P, Liao C, Tejera A, McNees C, Flores JM, Fernandez-Capetillo O, Tarsounas M, Blasco MA (2009) Increased telomere fragility and fusions resulting from TRF1 deficiency lead to degenerative pathologies and increased cancer in mice. *Genes Dev* **23:** 2060–2075

McGuinness BE, Hirota T, Kudo NR, Peters JM, Nasmyth K (2005) Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. *PLoS Biol* **3:** e86

Mendez J, Stillman B (2000) Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol Cell Biol* **20:** 8602–8612

Merkenschlager M (2010) Cohesin: a global player in chromosome biology with local ties to gene regulation. *Curr Opin Genet Dev* **20:** 555–561

Michaelis C, Ciosk R, Nasmyth K (1997) Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91:** 35–45

Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M (2009) Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO J* **28:** 1234–1245

Munoz P, Blanco R, Flores JM, Blasco MA (2005) XPF nuclease-dependent telomere loss and increased DNA damage in mice overexpressing TRF2 result in premature aging and cancer. *Nat Genet* **37:** 1063–1071

Naim V, Rosselli F (2009) The FANC pathway and BLM collaborate during mitosis to prevent micro-nucleation and chromosome abnormalities. *Nat Cell Biol* **11:** 761–768

Nasmyth K, Haering CH (2009) Cohesin: its roles and mechanisms. *Annu Rev Genet* **43:** 525–558

Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters JM, Murrell A (2009) Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet* **5:** e1000739

Nishiyama T, Ladurner R, Schmitz J, Kreidl E, Schleiffer A, Bhaskara V, Bando M, Shirahige K, Hyman AA, Mechtler K, Peters JM (2010) Sororin mediates sister chromatid cohesion by antagonizing Wapl. *Cell* **143:** 737–749

Norio P, Schildkraut CL (2001) Visualization of DNA replication on individual Epstein-Barr virus episomes. *Science* **294:** 2361–2364

Palidwor GA, Shcherbinin S, Huska MR, Rasko T, Stelzl U, Arumughan A, Foulle R, Porras P, Sanchez-Pulido L, Wanker EE, Andrade-Navarro MA (2009) Detection of alpha-rod protein repeats using a neural network and application to huntingtin. *PLoS Comput Biol* **5:** e1000304

Palm W, de Lange T (2008) How shelterin protects mammalian telomeres. *Annu Rev Genet* **42:** 301–334

Remeseiro S, Cuadrado A, Gómez-López G, Pisano DG, Losada A (2012) A unique role of cohesin-SA1 in gene regulation and development. *EMBO J* **31:** 2090–2102

Revenkova E, Eijpe M, Heyting C, Hodges CA, Hunt PA, Liebe B, Scherthan H, Jessberger R (2004) Cohesin SMC1 beta is required for meiotic chromosome dynamics, sister chromatid cohesion and DNA recombination. *Nat Cell Biol* **6:** 555–562

Rhodes JM, Bentley FK, Print CG, Dorsett D, Misulovin Z, Dickinson EJ, Crosier KE, Crosier PS, Horsfield JA (2010) Positive regulation of c-Myc by cohesin is direct, and evolutionarily conserved. *Dev Biol* **344:** 637–649

Rivera T, Losada A (2009) Shugoshin regulates cohesion by driving relocalization of PP2A in Xenopus extracts. *Chromosoma* **118:** 223–233

Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A (2008) CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci USA* **105:** 8309–8314

Salmon ED, Cimini D, Cameron LA, DeLuca JG (2005) Merotelic kinetochores in mammalian tissue cells. *Philos Trans R Soc Lond B Biol Sci* **360:** 553–568

Schnutgen F, De-Zolt S, Van Sloun P, Hollatz M, Floss T, Hansen J, Altschmied J, Seisenberger C, Ghyselinck NB, Ruiz P, Chambon P,

Wurst W, von Melchner H (2005) Genomewide production of multipurpose alleles for the functional analysis of the mouse genome. *Proc Natl Acad Sci USA* **102:** 7221–7226

Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, Marks H, Adams DJ, Schatz DG, Aragon L, Fisher AG, Krangel MS, Nasmyth K, Merkenschlager M (2011) A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* **476:** 467–471

Sfeir A, Kosiyatrakul ST, Hockemeyer D, MacRae SL, Karlseder J, Schildkraut CL, de Lange T (2009) Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell* **138:** 90–103

Smogorzewska A, Karlseder J, Holtgreve-Grez H, Jauch A, de Lange T (2002) DNA ligase IV-dependent NHEJ of deprotected mammalian telomeres in G1 and G2. *Curr Biol* **12:** 1635–1644

Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, Harris BT, Toretsky JA, Rosenberg SA, Shukla N, Ladanyi M, Samuels Y, James CD, Yu H, Kim JS, Waldman T (2011) Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science* **333:** 1039–1043

Sumara I, Vorlaufer E, Gieffers C, Peters BH, Peters JM (2000) Characterization of vertebrate cohesin complexes and their regulation in prophase. *J Cell Biol* **151:** 749–762

Torres-Rosell J, De Piccoli G, Cordon-Preciado V, Farmer S, Jarmuz A, Machin F, Pasero P, Lisby M, Haber JE, Aragon L (2007) Anaphase onset before complete DNA replication with intact checkpoint responses. *Science* **315:** 1411–1415

Toyoda Y, Yanagida M (2006) Coordinated requirements of human topo II and cohesin for metaphase centromere alignment under Mad2-dependent spindle checkpoint surveillance. *Mol Biol Cell* **17:** 2287–2302

Vagnarelli P, Morrison C, Dodson H, Sonoda E, Takeda S, Earnshaw WC (2004) Analysis of Scc1-deficient cells defines a key metaphase role of vertebrate cohesin in linking sister kinetochores. *EMBO Rep* **5:** 167–171

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451:** 796–801

Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, Housman DE, Amon A (2008) Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science* **322:** 703–709

Xu H, Balakrishnan K, Malaterre J, Beasley M, Yan Y, Essers J, Appeldoorn E, Tomaszewski JM, Vazquez M, Verschoor S, Lavin MF, Bertoncello I, Ramsay RG, McKay MJ (2010) Rad21-cohesin haploinsufficiency impedes DNA repair and enhances gastrointestinal radiosensitivity in mice. *PLoS ONE* **5:** e12112

Xu H, Beasley MD, Warren WD, van der Horst GT, McKay MJ (2005) Absence of mouse REC8 cohesin promotes synapsis of sister chromatids in meiosis. *Dev Cell* **8:** 949–961

Xu H, Tomaszewski JM, McKay MJ (2011) Can corruption of chromosome cohesion create a conduit to cancer? *Nat Rev Cancer* **11:** 199–210

Ye J, Lenain C, Bauwens S, Rizzo A, Saint-Leger A, Poulet A, Benarroch D, Magdinier F, Morere J, Amiard S, Verhoeyen E, Britton S, Calsou P, Salles B, Bizard A, Nadal M, Salvati E, Sabatier L, Wu Y, Biroccio A *et al* (2010) TRF2 and apollo cooperate with topoisomerase 2alpha to protect human telomeres from replicative damage. *Cell* **142:** 230–242

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada. **Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.** The EMBO Journal (2012) 31, 2076–2089.

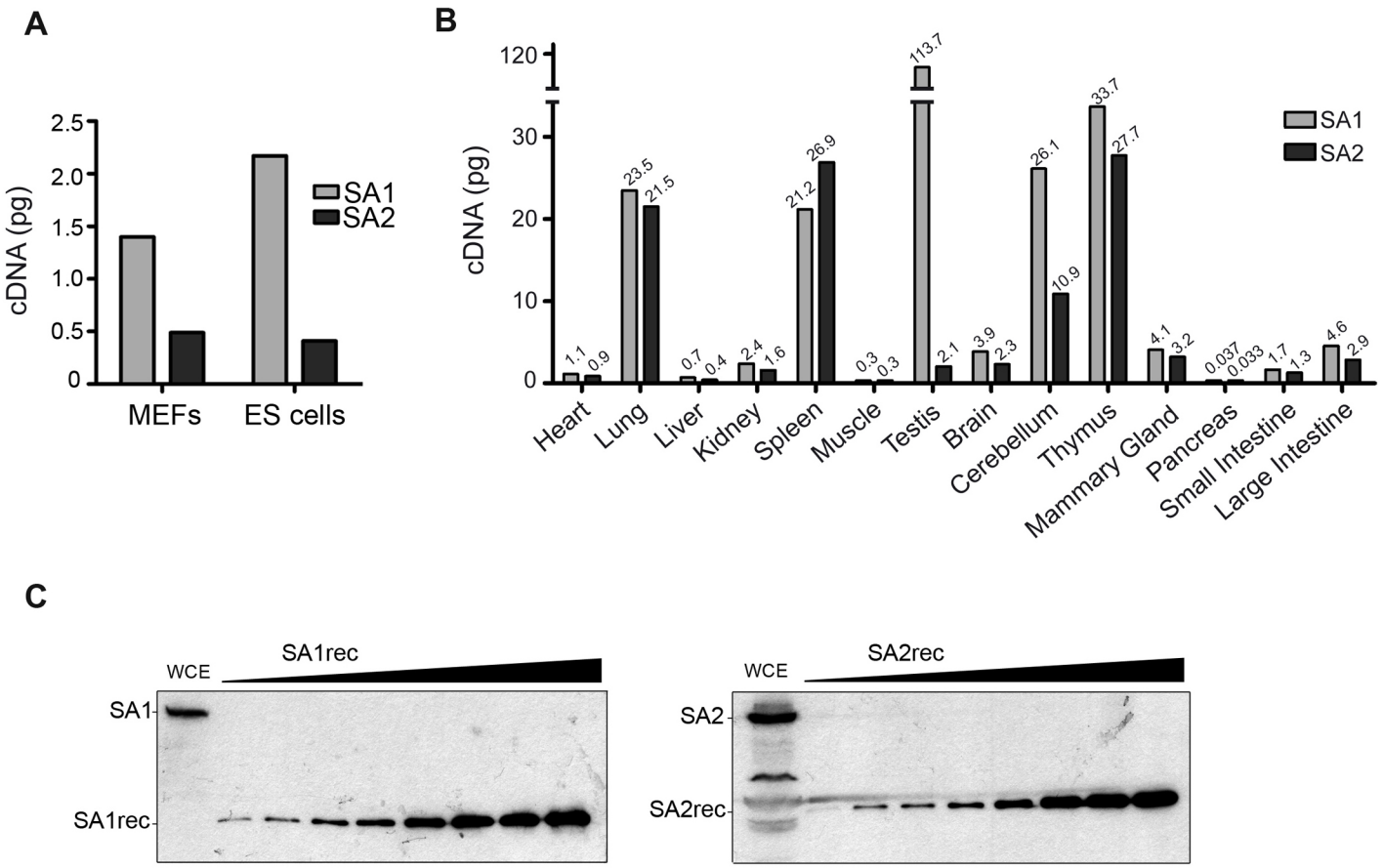


Figure S1. SA1 and SA2 are ubiquitously expressed in mouse tissues.

To address the relative abundance of cohesin complexes containing either SA1 or SA2 in mouse cells, we developed a protocol to quantify the absolute amount of each mRNA by quantitative RT-PCR. We found that the SA1 transcripts are 1 to 3 times more abundant than those of SA2 in (**A**) MEFs, ES cells and (**B**) in cells from different mouse tissues. In testis, the difference is much higher, suggesting that SA1 may have an important function in meiosis. (**C**) An aliquot of MEFs whole cell extract (WCE) was immunoblotted along with increasing amounts of recombinant SA1 or SA2 proteins (0.25 to 8 ng). Comparison and quantitation of the signals obtained indicate that SA2 is 3 times more abundant than SA1 in MEFs.

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada. **Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.** The EMBO Journal (2012) 31, 2076–2089.

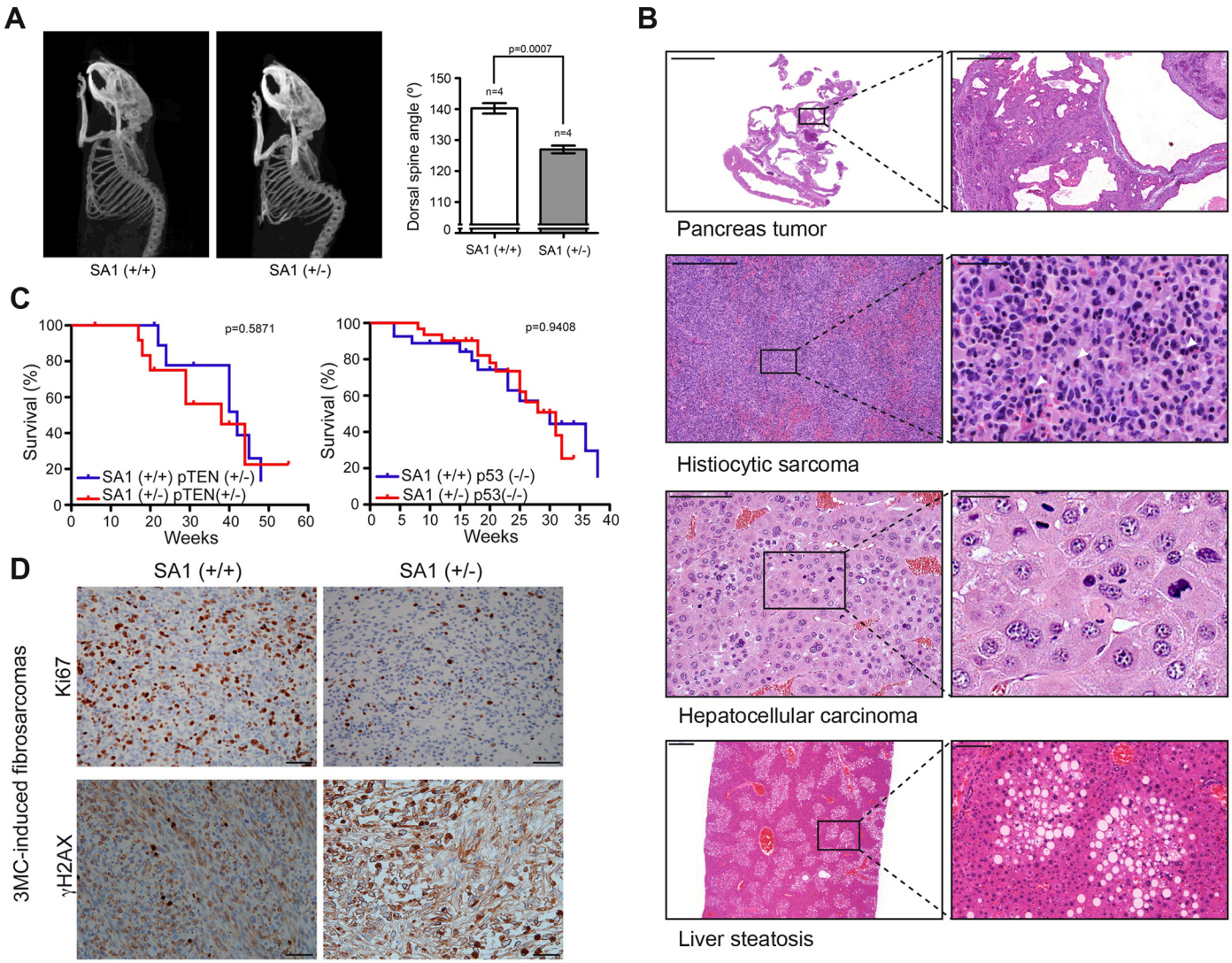


Figure S2. Reduced lifespan, increased incidence of spontaneous tumors, but higher resistance against chemically induced tumors in SA1 heterozygous mice.

(A) Acute kyphosis a sign of premature ageing, is observed in SA1 heterozygous mice by Computed Tomography. Kyphosis is evidenced by a reduced dorsal spine angle measured as the angle formed by T12, L3 and L5 vertebrae. Four 24-month old animals of each genotype were analyzed.

(B) Hematoxylin-eosin staining of tissue sections from SA1 heterozygous mice whose pathology was analyzed. From top to bottom: a very aggressive pancreas tumor that resembles the human pancreatic IPMN (scale bars, 5000 and 500 μm, left and right respectively); histiocytic sarcoma (arrowheads point to aberrant mitosis) (scale bars, 500 and 50 μm); hepatocellular carcinoma (scale bars, 100 and 20 μm); liver steatosis, a non-tumoral pathology extremely frequent in SA1 heterozygous mice that might end up in hepatocellular carcinomas (scale bars, 500 and 100 μm).

(C) Kaplan-Meier survival curves for SA1 heterozygous and wildtype mice in either p53-null background (n=31 and 27, respectively, left) or PTEN heterozygous background (n=14 and 13, respectively, right). No differences were found.

(D) Immunohistochemical detection of Ki67 and γH2AX on tissue sections from the fibrosarcomas induced by 3-MC in mice of the indicated genotype (quantification is shown in Fig. 2D).

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada. **Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.** The EMBO Journal (2012) 31, 2076–2089.

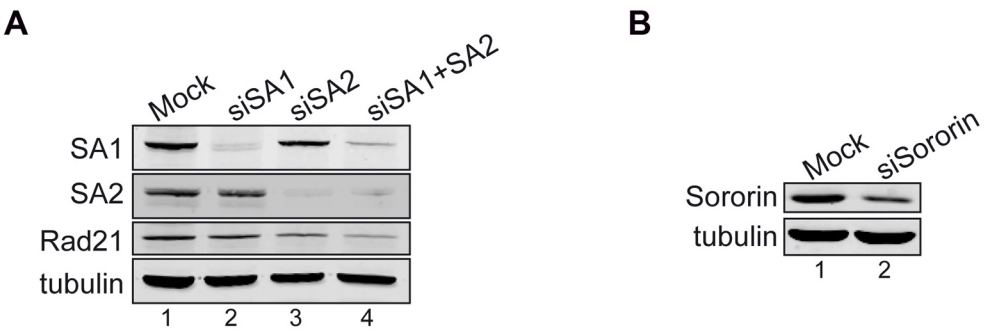


Figure S3. Downregulation of cohesin subunits and Sororin in C2C12 cells.

(A) Immunoblot analysis of cell extracts of mouse C2C12 cells treated with siRNAs against **(A)** SA1 (siSA1), SA2 (siSA2) or both (siSA1+SA2) or **(B)** Sororin (siSororin) or with no siRNA (mock). Tubulin is used as loading control.

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada. **Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.** The EMBO Journal (2012) 31, 2076–2089.

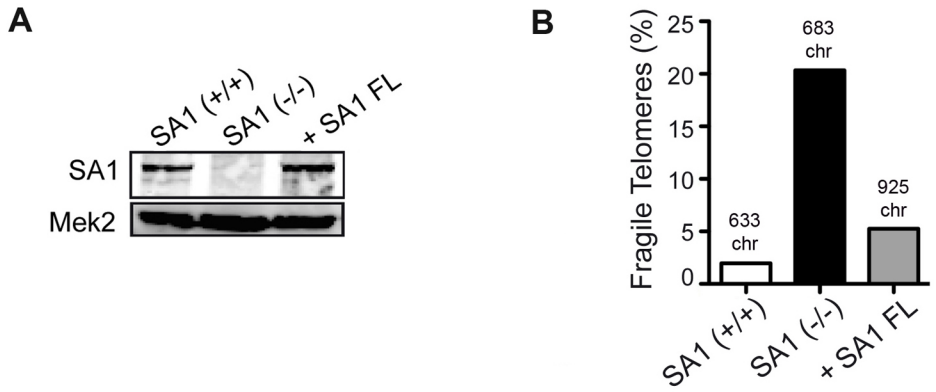


Figure S4. Expression of full-length SA1 in SA1-null MEFs rescues telomere fragility.

(A) Immunoblot analysis of whole cell extracts from wildtype, SA1-null cells and SA1-null cells transfected with full-length SA1 (+ SA1 FL). Mek2 is used as loading control.

(B) Percentage of fragile telomeres observed in the conditions mentioned above. Expression of full-length SA1 protein restores substantially the telomere structure.

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada. **Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.** The EMBO Journal (2012) 31, 2076–2089.

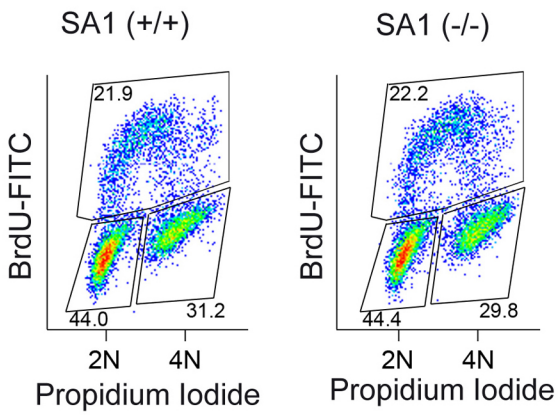
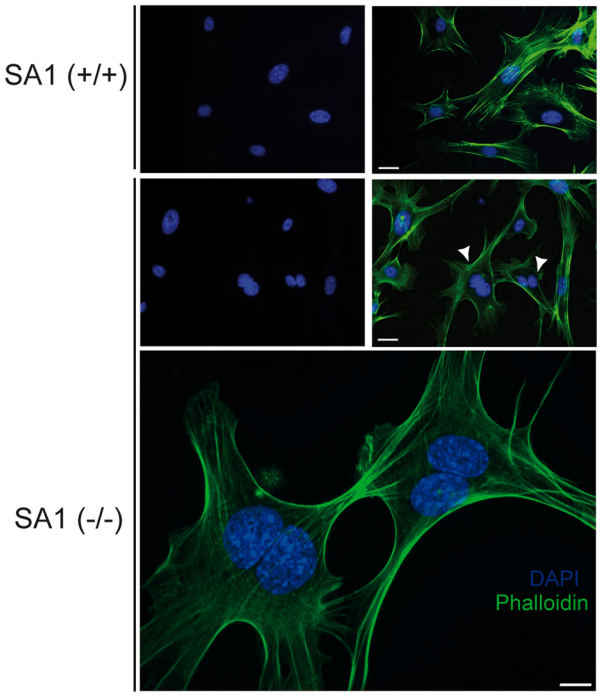


Figure S5. Cell cycle profiling of cell populations analyzed by SMARD.

Cell cycle profiles and BrdU incoporation of wildtype and SA1-null MEFs at the time point collected for SMARD analysis suggest that there is not a general defect in S phase progression.

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada. **Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.** The EMBO Journal (2012) 31, 2076–2089.



	Binucleated cells (%)
SA1 (+/+)	5.6
SA1 (-/-)	29.1

Figure S6. Presence of binucleated cells in fixed SA1-null cultures.

Quantification of the percentage of binucleated cells found in cultures of wildtype and SA1-null MEFs stained with DAPI (blue) and phalloidin (green). (n≥250 cells from two different clones per genotype). Examples of mononucleated and binucleated MEFs (arrowheads) are shown. Scale bars, 40 μm (top and middle panels) and 10 μm (bottom panel).

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada. **Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.** The EMBO Journal (2012) 31, 2076–2089.

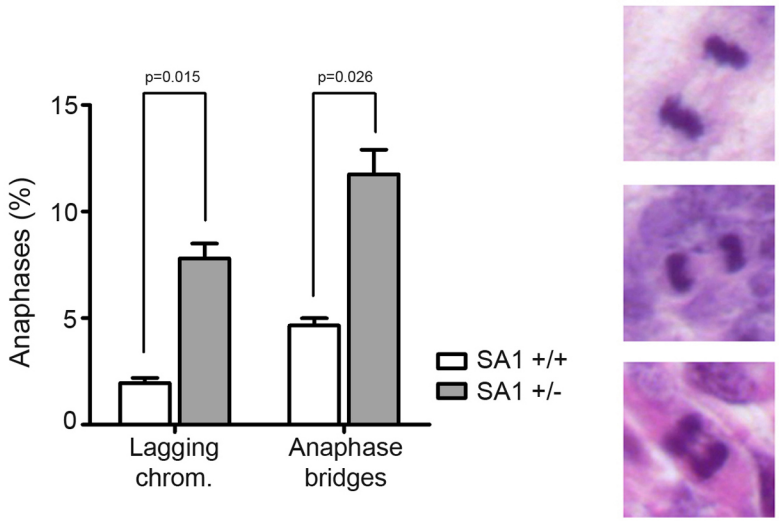


Figure S7. Chromosome segregation defects in SA1 (+/-) skin sections.

Frequency of anaphases with lagging chromosomes or anaphase bridges in HE-stained tissue sections from wildtype and SA1 heterozygous embryos. Examples of a normal anaphase (top), lagging chromosomes (middle) and anaphases bridges (bottom) are shown in the panels. (n≥50 cells per embryo from 2 independent embryos per genotype).

Silvia Remeseiro, Ana Cuadrado, María Carretero, Paula Martínez, William C Drosopoulos, Marta Cañamero, Carl L Schildkraut, María A Blasco and Ana Losada. **Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres.** The EMBO Journal (2012) 31, 2076–2089.

Observed frequencies SA1 (+/-) X SA1 (+/-)

E	(+/+)	(+/-)	(-/-)	Total
11.5	26.2	50.0	23.8	42
12.5	26.6	58.9	14.6	192
13.5	47.2	36.1	16.7	36
14.5	34.4	57.8	7.8	64
15.5	55.6	44.4	0.0	9
16.5	62.5	37.5	0.0	16
17.5	30.9	56.1	13.0	123
18.5	36.0	54.0	10.0	50
Newborn	41.9	58.1	0	>300

A unique role of cohesin-SA1 in gene regulation and development

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada.

The EMBO Journal (2012) 31, 2090–2102

In this study, taking advantage of the SA1-null model as a tool, we wanted to further explore the functional specificity of cohesin-SA1 and cohesin-SA2. In order to gain insight into this, we first compared their genome-wide distributions. We report, for the first time, the genome-wide binding sites of four different cohesin subunits in both wildtype and SA1-null cells from murine origin. We found that cohesin-SA1 is enriched at promoters and that the regions containing the highest density of cohesin-SA1 molecules are mainly promoters. More importantly we demonstrate that cohesin-SA1 determines the distribution of cohesin. In the absence of SA1, cohesin is significantly less abundant at gene promoters and relocates to genomic positions featured by low cohesin occupancy and absence of CTCF. This suggests that SA2 is incapable of replacing SA1, particularly at these genomic regions. As a consequence, lack of SA1 alters gene expression, affecting genes involved in developmental pathways relevant to CdLS. We also show that the presence of SA1 at the promoters of certain genes positively regulates their expression (e.g. myc, Protocadherins), a function that cannot be assumed by SA2. Other cohesin-SA1 binding sites outside promoters are relevant for proper gene expression. Indeed, the absence of SA1 also alters the cohesin-binding pattern along some gene clusters leading to dysregulation of the genes within. All this, together with the fact that SA1-null embryos present features reminiscent of CdLS, led us to hypothesize that impaired function of cohesin-SA1 in regulation of gene expression may underlie the etiology of Cornelia de Lange syndrome.

In this project I have worked in team with Ana Cuadrado, coauthor of the paper, and my contribution comprises the design, performance, analysis and interpretation of all the experiments, as well as writing of the manuscript, under the supervision of the thesis director Dr. Ana Losada.

The EMBO Journal (2012) 31, 2090–2102 | © 2012 European Molecular Biology Organization | All Rights Reserved 0261-4189/12
www.embojournal.org

THE EMBO JOURNAL

A unique role of cohesin-SA1 in gene regulation and development

Silvia Remeseiro[1,3], Ana Cuadrado[1,3], Gonzalo Gómez-López[2], David G Pisano[2] and Ana Losada[1,*]

[1]Chromosome Dynamics Group, Molecular Oncology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain and [2]Bioinformatics Unit, Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Vertebrates have two cohesin complexes that consist of Smc1, Smc3, Rad21/Scc1 and either SA1 or SA2, but their functional specificity is unclear. Mouse embryos lacking SA1 show developmental delay and die before birth. Comparison of the genome-wide distribution of cohesin in wild-type and SA1-null cells reveals that SA1 is largely responsible for cohesin accumulation at promoters and at sites bound by the insulator protein CTCF. As a consequence, ablation of SA1 alters transcription of genes involved in biological processes related to Cornelia de Lange syndrome (CdLS), a genetic disorder linked to dysfunction of cohesin. We show that the presence of cohesin-SA1 at the promoter of myc and of protocadherin genes positively regulates their expression, a task that cannot be assumed by cohesin-SA2. Lack of SA1 also alters cohesin-binding pattern along some gene clusters and leads to dysregulation of genes within. We hypothesize that impaired cohesin-SA1 function in gene expression underlies the molecular aetiology of CdLS.

The EMBO Journal (2012) **31,** 2090–2102. doi:10.1038/emboj.2012.60; Published online 13 March 2012
Subject Categories: chromatin & transcription; cell cycle
Keywords: CdLS; ChIP-sequencing; embryonic development; mouse model; transcription

Introduction

In addition to its role in sister chromatid cohesion, essential for accurate chromosome segregation and postreplicative DNA repair (Losada and Hirano, 2005; Peters *et al*, 2008; Nasmyth and Haering, 2009), cohesin contributes to transcriptional regulation (Rollins *et al*, 2004; Dorsett *et al*, 2005; Horsfield *et al*, 2007; Hallson *et al*, 2008; Pauli *et al*, 2008, 2010; Schaaf *et al*, 2009; Lin *et al*, 2011), DNA replication (Guillou *et al*, 2010), V(D)J recombination (Degner *et al*, 2009) and T-cell receptor rearrangement (Seitan *et al*, 2011). Chromosome conformation capture (3-C) studies and other experimental evidences suggest that cohesin performs these functions by promoting the formation of chromatin loops in collaboration with additional factors (Hadjur *et al*, 2009; Mishiro *et al*, 2009; Nativio *et al*, 2009; Hou *et al*, 2010; Chien *et al*, 2011; Seitan *et al*, 2011). One such factor is CTCF, which colocalizes with cohesin at many sites along the human and mouse genomes (Parelho *et al*, 2008; Rubio *et al*, 2008; Wendt *et al*, 2008). Cohesin is also found at non-CTCF-binding sites occupied by tissue-specific transcription factors (TFs), and possibly contributes to set up tissue-specific transcriptional programs (Schmidt *et al*, 2010). In murine embryonic stem cells, cohesin and the transcriptional coactivator Mediator facilitate DNA looping between the enhancers and promoters of some genes required to maintain pluripotency (Kagey *et al*, 2010).

The involvement of cohesin in all these different processes that are critical for cell proliferation and differentiation makes it difficult to determine which is most relevant for cohesinopathies, human syndromes caused by mutations in proteins related to cohesin (Bose and Gerton, 2010). The most prevalent cohesinopathy described to date is the Cornelia de Lange syndrome (CdLS), which is caused by heterozygous mutations in the cohesin loader Nipbl (Krantz *et al*, 2004; Tonkin *et al*, 2004) or, in a reduced number of cases, in the cohesin subunits Smc1 or Smc3 (Musio *et al*, 2006; Deardorff *et al*, 2007). CdLS is characterized by mental retardation, reduced body size, dysmorphic face, upper limb defects and several additional organ abnormalities (Liu and Krantz, 2009). Cells from patients rarely show cohesion defects and instead, microarray studies reveal altered patterns of gene expression (Castronovo *et al*, 2009; Liu *et al*, 2009). Nipbl-heterozygous mice exhibit many features of CdLS and display modest but significant transcriptional changes in mouse embryonic fibroblasts (MEFs) and embryonic brains (Kawauchi *et al*, 2009). Whether the wide range of defects observed in CdLS patients are elicited by expression changes in a few critical genes or by the sum of multiple small alterations is unclear (Muto *et al*, 2011). Moreover, the exact role of cohesin in CdLS pathogenesis is still far from understood.

Importantly, somatic vertebrate cells have two distinct versions of cohesin that consist of Smc1, Smc3, Rad21/Scc1 and either SA1 or SA2 (Losada *et al*, 2000; Sumara *et al*, 2000). The two SA proteins show 75% sequence identity along their central region and only differ in short regions at both ends of each protein. Cohesin-SA1 and cohesin-SA2 present similar cell-cycle regulation in terms of loading and dissociation from chromatin (Losada *et al*, 2000) and are found in all mouse tissues (Remeseiro *et al*, 2012). However, they are not functionally equivalent. Mice homozygous for a gene-trap allele that abolishes SA1 expression are embryonic lethal (Remeseiro *et al*, 2012). We have shown that cohesin-SA1 plays a specific role in telomere cohesion that is essential for efficient telomere replication. In contrast, cohesion at centromeres, which is critical for chromosome segregation, relies on cohesin-SA2 (Canudas and Smith, 2009; Solomon

*Corresponding author. Molecular Oncology Programme, Spanish National Cancer Research Centre (CNIO), Melchor Fernández Almagro 3, Madrid 28029, Spain. Tel.: +34 917328000/ext. 3470; Fax: +34 917328033; E-mail: alosada@cnio.es
[3]These two authors contributed equally to this work

Received: 29 September 2011; accepted: 20 February 2012; published online 13 March 2012

et al, 2011). Despite their robust centromeric cohesion, chromosome segregation is defective in SA1-deficient cells and often leads to aneuploidy, increasing the incidence of spontaneous tumours in heterozygous animals (Remeseiro *et al*, 2012).

To further understand the functional specificity of cohesin-SA1 and cohesin-SA2, we have compared their genomic distributions. We report, for the first time, the genome-wide binding sites of four different cohesin subunits in both wild-type and SA1-null mouse cells. We show that cohesin-SA1 is enriched at promoters and, more importantly, that it determines the distribution of cohesin. In the absence of SA1, cohesin is significantly less abundant at gene promoters and relocates to genomic positions featured by low cohesin occupancy and absence of CTCF. As a result, gene expression is altered, affecting genes involved in developmental pathways relevant to CdLS.

Results

SA1 is essential for embryonic development

We recently generated a genetically modified mouse model lacking the expression of Stag1 gene encoding the SA1 cohesin subunit (Remeseiro *et al*, 2012). Viability of SA1-null embryos strongly decreases by E12.5, but some exceptionally survive to E18.5 (see Supplementary Table S1 in Remeseiro *et al*, 2012). Those late embryos present a clear growth delay and general hypoplasia (Figure 1A). Their skin is much thinner, with less hair follicles and a reduced muscle layer compared with their wild-type littermates (Figure 1B, panel I). The developmental delay is also observed in organs such as kidney and liver (Figure 1B, panels II and III, respectively). Interestingly, SA1-null embryos present some features characteristic of CdLS in addition to the reduced body size. These include impairment of lipid metabolism, as illustrated by a dramatically diminished interscapular layer of brown adipose tissue (BAT) (Figure 1B, panel IV), and severe abnormalities in skeletal and bone development, featured by delayed ossification and defective calcium deposition (Figure 1B, panels V–VII). Both cohesin-SA1 and cohesin-SA2 are ubiquitously expressed in wild-type E17.5 embryos and,

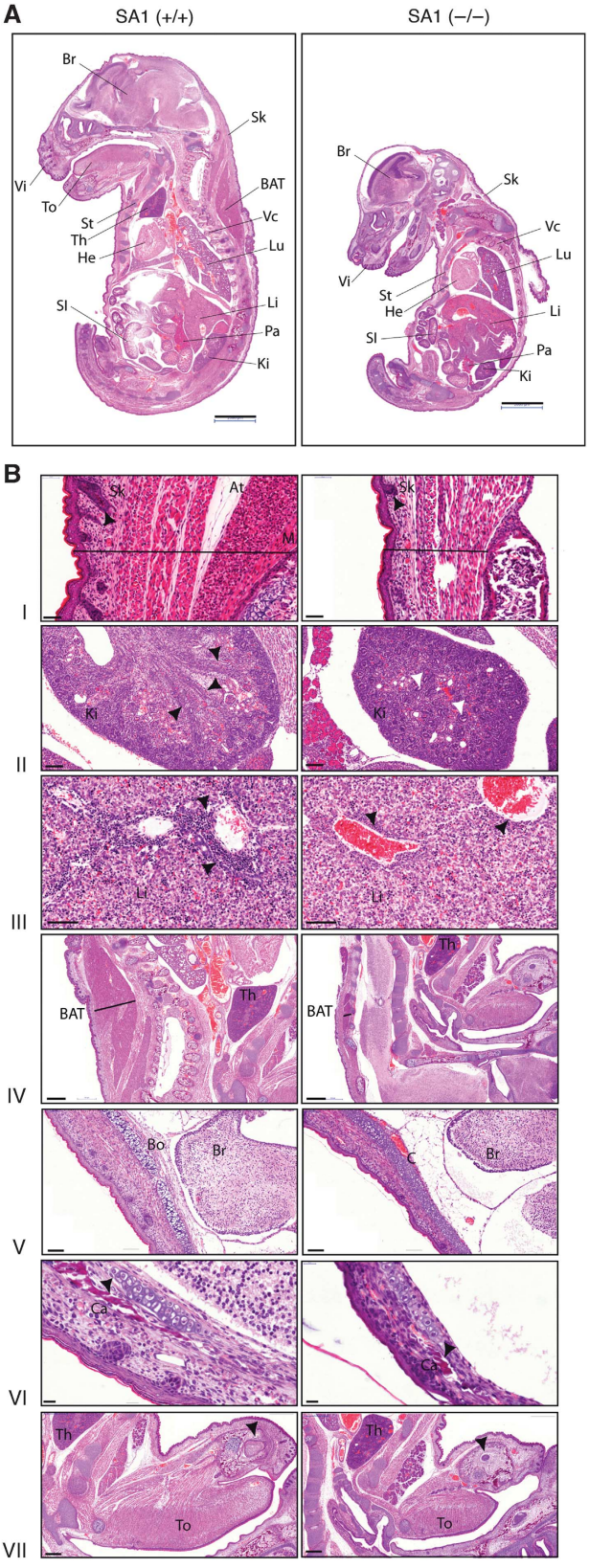


Figure 1 SA1-null embryos present growth delay. Histological analysis of HE-stained tissue sections from E17.5 embryos shows general hypoplasia and developmental delay in SA1-null embryos. (**A**) SA1-null embryos present a much reduced body size compared with their wild-type littermates. The following organs/tissues are indicated: BAT, brown adipose tissue; Br, brain; He, heart; Ki, kidney; Li, liver; Lu, lung; Pa, pancreas; SI, small intestine; Sk, skin; St, sternum; Th, thymus; To, tongue; Vc, vertebral column; Vi, vibrissa. Scale bars, 2 mm. (**B**) Detailed sections of different tissues showing either developmental delay or CdLS features. (I) Thinner skin (black line) in SA1-null embryos with less hair follicles (arrowheads), reduced adipose tissue (At) and thinner muscle layer (M). Scale bars, 50 μm. (II) Kidney sections showing collecting ducts already formed in wild-type (black arrowheads) and not radially arranged in SA1-null embryos. In the latter, glomeruli are dispersed all over as primitive glomeruli at earlier stages (white arrowheads). Scale bars, 100 μm. (III) Decreased number of haematopoietic precursors in livers from SA1-null embryos compared with wild-type (arrowheads). Scale bars, 100 μm. (IV) Thickness of the interscapular layer of BAT is dramatically diminished in SA1-null embryos (black line). Scale bars, 500 μm. (V) Delayed intramembranous ossification in the cranium in the absence of SA1. Notice the presence of cartilage (C) in SA1-null embryos instead of bone (Bo) already formed in the wild-type. Scale bars, 100 μm. (VI) Defective bone calcification (Ca) in the cranium from SA1-null embryos. Scale bars, 20 μm. (VII) Delay in dentition in SA1-null embryos. Arrowheads in both sections indicate tooth buds. Scale bars, 500 μm.

importantly, SA2 expression is not altered in SA1-null embryonic tissues, including those described above (Supplementary Figure S1A). Protein and mRNA levels of other cohesin subunits remain invariable in the absence of SA1 (shown for brain in Supplementary Figure S1B and C). Moreover, the amount of total cohesin bound to chromatin does not change in MEFs lacking SA1 (Figure 4G in Remeseiro *et al*, 2012). Thus, it is unlikely that the inability of cohesin-SA2 to compensate the cohesin-SA1 deficiency has merely a quantitative basis. Instead, cohesin-SA1 must perform specific function(s) in embryonic development that cannot be assumed by cohesin-SA2.

Genome-wide distribution of cohesin-SA1 and cohesin-SA2

In order to assess if a differential distribution of the two cohesin complexes contributes to their functional specificity, we performed chromatin immunoprecipitation followed by massive parallel DNA sequencing (ChIP-seq) in primary MEFs (E12.5) using SA1, SA2, SMC1 and SMC3-specific antibodies (Supplementary Figure S2A and B). Genome-wide analysis of the sequenced tags using MACS peak detection algorithm (see Materials and methods) defined 25 737 binding sites for SA1, 7741 for SA2, 23 994 for SMC1 and 15 546 for SMC3, with a cutoff *P*-value of 10^{-5} and FDR < 0.1 (Figure 2A and B; ChIP-seq data summary in Supplementary Tables S1 and S2). A subset of these regions was further validated by quantitative PCR analysis (ChIP-qPCR; Supplementary Figure S2C, D and E). As expected for proteins forming a complex, a large overlap was observed between the binding sites obtained with SMC1 and SMC3 antibodies, as well as between those seen with SMC1 and either SA1 or SA2 antibodies (Figure 2C). Lack of a more complete overlap is likely attributable to slight differences in the accessibility of the antibodies to the complex depending on its chromatin environment. Since cohesin has been reported to colocalize with the insulator protein CTCF, we compared all our cohesin data sets with a data set of CTCF-binding sites from a murine cell line available online (ENCODE/Stanford/Yale data set). Most positions detected with any given cohesin antibody correspond to a CTCF-binding site, with the SA2 data set showing the lowest overlap (63%; Supplementary Figure S3A). Thus, our ChIP-seq data, obtained with four different antibodies, provide an accurate map of cohesin-binding sites along the mouse genome.

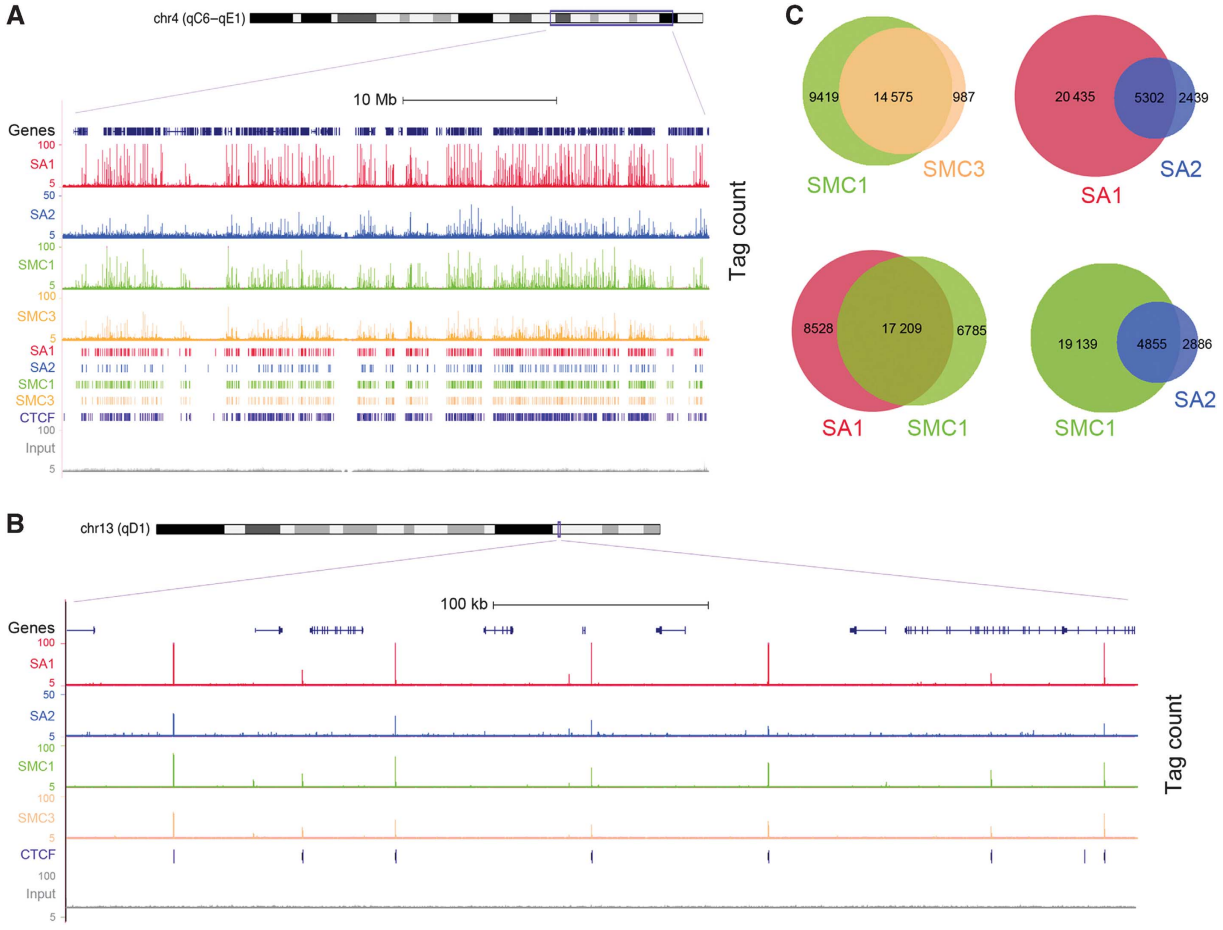


Figure 2 Genome-wide distribution of cohesin-SA1 and cohesin-SA2 in the mouse genome. (**A**) Representative genomic distribution of SA1, SA2, SMC1 and SMC3 in a region of mouse chromosome 4. Tag counts (upper) and called peak tracks (lower) are shown. A track with CTCF-binding sites (from ENCODE/Stanford/Yale data set) and the input are also depicted. Higher magnification of a region of chromosome 13 is shown in (**B**). (**C**) Venn diagrams showing the overlapping between the binding sites identified in wild-type MEFs by ChIP-seq with the indicated antibodies.

Cohesin distribution and enrichment at promoters depend on SA1

Cohesin-SA1 and cohesin-SA2 binding sites are distributed among gene-associated and intergenic regions to a similar extent in wild-type MEFs (Figure 3A, left and middle pie charts). Importantly, cohesin-SA1 is highly enriched at regions 1-kilobase (kb) upstream transcription start sites (TSS) considering that they represent only 0.62% of the mouse genome (Figure 3B, left). The enrichment is significantly less noticeable for cohesin-SA2. This difference is even more remarkable if we compare enrichment at TSS of cohesin-binding sites that are exclusive for SA1 or SA2 (Figure 3B, right). In addition, the frequency of binding at TSS is higher for cohesin-SA1 than for cohesin-SA2 (Figure 3C, left). We next asked whether the absence of SA1 affects cohesin distribution. Indeed, ChIP-seq analysis of SA1-null MEFs with the SA2 antibody resulted in 6349 peaks (*P*-value of 10^{-5}, FDR < 0.1) from which only 45% overlap with the SA2 peaks found in wild-type MEFs (Supplementary Figure S3B; Supplementary Tables S1 and S2). The new positions occupied by SA2 correspond to binding sites preferentially located at intergenic regions (Figure 3A, pie chart on the right). The frequency of SA2 binding at TSS is further reduced in SA1-null cells, which suggests that SA2 is unable to replace SA1 particularly at these genomic positions (Figure 3C, right). Interestingly, we observed a gradual increase in frequency of binding at TSS as the cohesin-bound regions lengthen only in the case of cohesin-SA1, but not for cohesin-SA2 in either wild-type or SA1-null cells (Figure 3D). This indicates that genomic regions containing the highest density of cohesin-SA1 molecules are mainly gene promoters.

To further characterize the redistribution of cohesin-SA2 in the absence of SA1, we performed ChIP-seq with SMC1 and SMC3 antibodies in SA1-null MEFs. The number of cohesin positions detected with these antibodies in the SA1-null MEFs (46 997 and 34 800, respectively) is higher than the number detected with the SA2 antibody, most likely due to a more limited sensibility of the latter (Supplementary Tables S1 and S2). The fact that the number of peaks doubles in both cases with respect to wild-type MEFs suggests a major redistribution of cohesin in the absence of SA1 (Supplementary Figure S3C). Of all SMC1-binding sites identified in SA1-null MEFs, 45% correspond to cohesin-SA1 positions in wild-type cells (group 'a' in Figure 3E). The remaining 55% (group 'b') show significantly lower peak intensity than those in group 'a', and the same is true for the SMC3-binding sites in SA1-null MEFs (groups 'c' and 'd' in Figure 3E; examples are shown in Figure 3F, left). Binding to some of these regions with lower cohesin occupancy was validated by ChIP-qPCR (Supplementary Figure S2C, lower panels). These positions have an additional feature. Whereas most sites bound by SMC1 and SMC3 in wild-type MEFs are also CTCF sites (71 and 82% respectively, see charts in Figure 3F, right), the overlap decreases to 46 and 58% in SA1-null cells, mostly due to the modest overlapping of the new positions with CTCF (groups 'b' and 'd', 19 and 27%, respectively). These results indicate that in the absence of SA1, cohesin spreads to CTCF-negative positions in which the complex shows significantly weaker enrichment compared with the positions defined in the wild-type cells. Taken all together, we conclude that cohesin accumulation at promoters and CTCF sites depends largely on the presence of SA1. This points to a unique role of cohesin-SA1 in transcriptional regulation.

Transcriptome changes in SA1-null cells affect CdLS-related functions

To obtain a global view of expression changes in the absence of SA1, we analysed the transcriptomes of SA1-null and wild-type MEFs using microarrays (Supplementary Table S3). An enrichment analysis of Gene Ontology (GO) terms revealed profound functional differences between both expression profiles. Transcriptional changes in the SA1-null cells led to upregulation of GO-defined processes related to phagocytosis, endocytosis and apoptosis. In contrast, processes associated with abnormalities observed in CdLS patients, such as limb and skeletal system morphogenesis, heart and lung development and lipid metabolism, were downregulated (Figure 4A; Supplementary Table S4). Moreover, we found a significant overlap (GO Comparative Analysis, FDR < 0.05) between the transcriptional changes linked to the loss of cohesin-SA1 and those reported for Nipbl-heterozygous MEFs (Kawauchi *et al*, 2009) (Figure 4B; Supplementary Table S5). A stringent statistical analysis (FDR < 0.15) identified 55 differentially expressed genes (DEGs), many of them significantly upregulated in the SA1-null MEFs (Supplementary Table S6). Validation of microarray data was performed by qPCR for a subset of 15 genes (Supplementary Figure S4). To determine if this transcriptional regulation is SA1 specific, we knocked down either SA1 or SA2 in MEFs by siRNA. Transient downregulation of SA1 significantly reduced the expression of 11 out of 14 genes selected from those found to be regulated by SA1 in the microarray analysis (Figure 4C). In contrast, reduction of SA2 expression only affected two of the tested genes, supporting the specificity of cohesin-SA1 in this regulation. Since SA1 is present in the vicinity of the TSS in only six of the DEGs (Supplementary Table S6), it is likely that cohesin-SA1 affects gene expression by additional mechanisms that do not require its presence at promoters. According to recent literature, cohesin promotes the proper topological organization of gene expression domains or gene clusters. Consistent with this role for cohesin-SA1, we observed that 18 out of 55 DEGs are located in close proximity of one another (< 800 kb; Figure 4D). Finally, cohesin may indirectly affect the expression of a gene by regulating a factor required for its transcription (Horsfield *et al*, 2007). As described in the next sections, we have found evidence of the importance of cohesin-SA1 in these different modes of regulation of gene expression by cohesin.

Cohesin-SA1 positively regulates c-myc expression

Myc is a major modulator of cell proliferation, growth and differentiation whose transcription is regulated by cohesin in different organisms (Stedman *et al*, 2008; Rhodes *et al*, 2010). Mutant mice expressing reduced levels of Myc have decreased body mass owing to multiorgan hypoplasia (Trumpp *et al*, 2001). Our ChIP-seq data show that the c-myc gene constitutes one of the most prominent SA1-binding regions in the mouse genome (Figure 5A). Quantitative qPCR analyses in MEFs validate this result and show that SMC1 levels in this region are severely reduced in the SA1-null cells (Supplementary Figure S2D). Expression of myc target genes is significantly altered in these cells (FDR = 0.06; Supplementary Figure S5), supporting the rele-

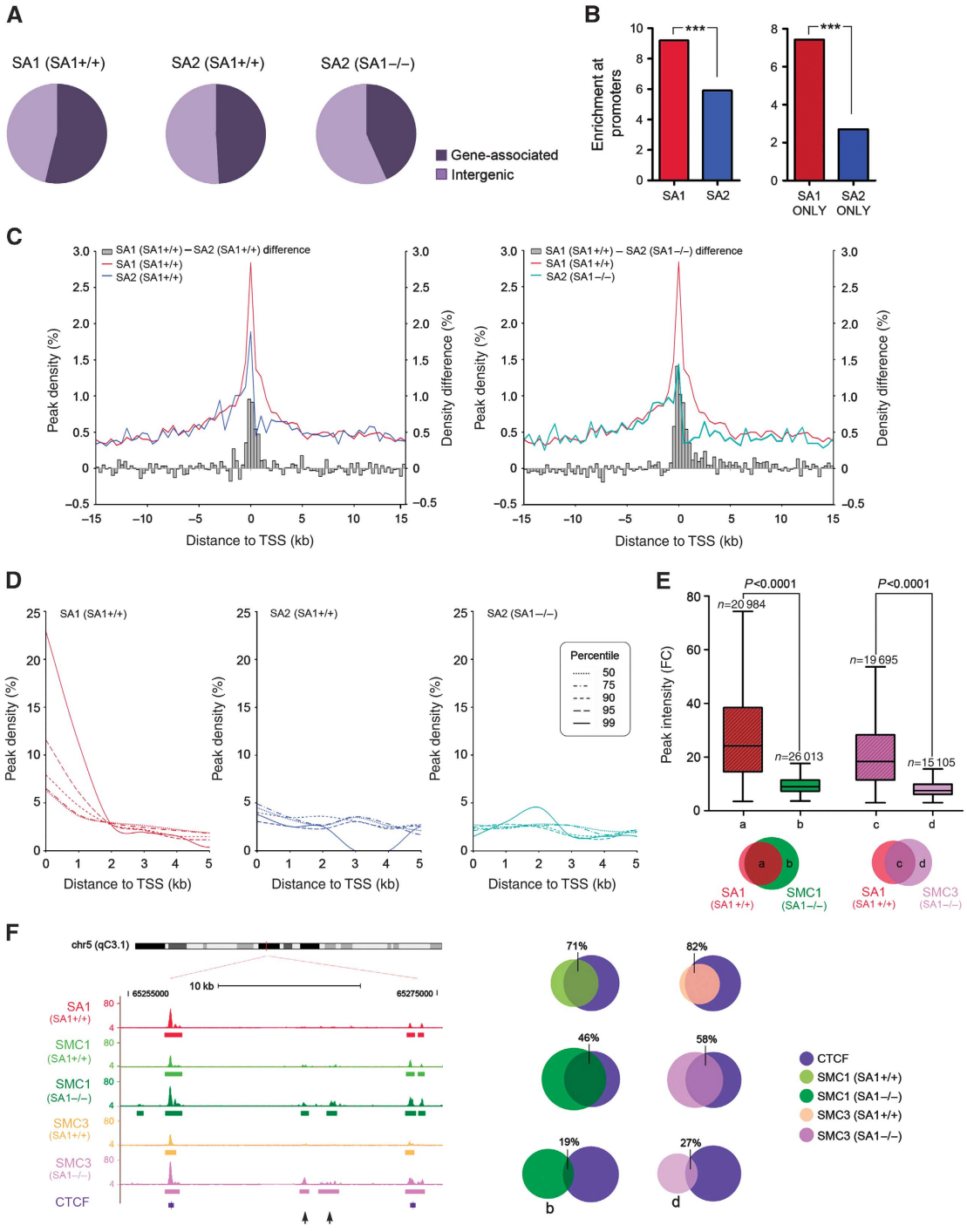


Figure 3 Cohesin distribution depends on SA1. (**A**) Distribution of SA1 and SA2 in SA1+/+ (wild-type) cells and of SA2 in SA1−/− cells represented as percentage of peaks detected at gene-associated regions (including 1 kb upstream TSS, Gene Body and 1 kb downstream TTS) and intergenic regions. Notice that SA2 relocates towards intergenic regions in the absence of SA1. (**B**) Left: The percentage of SA1 and SA2 binding sites located in regions 1 kb upstream TSS in wild-type cells was normalized against the frequency of these regions in the genome and displayed as fold enrichment. Right: Same analysis performed for positions that are exclusive for SA1 or SA2 (shadowed bars). ***$P<0.0001$. (**C**) Cohesin distribution around TSS (± 15 kb) defined as peak density (%). The histogram represents the difference in peak density between SA1 and SA2. (**D**) Frequency of binding at TSS in percentiles of peak length distribution (P50, P75, P90, P95 and P99). (**E**) Peak intensity (measured as fold change) for SMC1 and SMC3 peaks identified in SA1-null MEFs that overlap (groups 'a' and 'c') or not (groups 'b' and 'd') with SA1 peaks, as indicated in the Venn diagrams below the graph. Notice that cohesin positions in groups 'a' and 'c' present significantly higher occupancy (medians = 24.4 and 18.5, respectively) than those in groups 'b' and 'd' (medians = 9.2 and 7.6, respectively). See Supplementary Figure S3B for the number of peaks in each group. (**F**) Left: Representative image showing the redistribution of SMC1 and SMC3 in SA1-null cells towards low occupancy sites (arrowheads). Right: Venn diagrams showing the overlapping between CTCF-binding sites and the cohesin-binding sites indicated, including those defined as groups 'b' and 'd' in (**E**). More detailed information regarding the peak number within each group is shown in Supplementary Figure S3A and B.

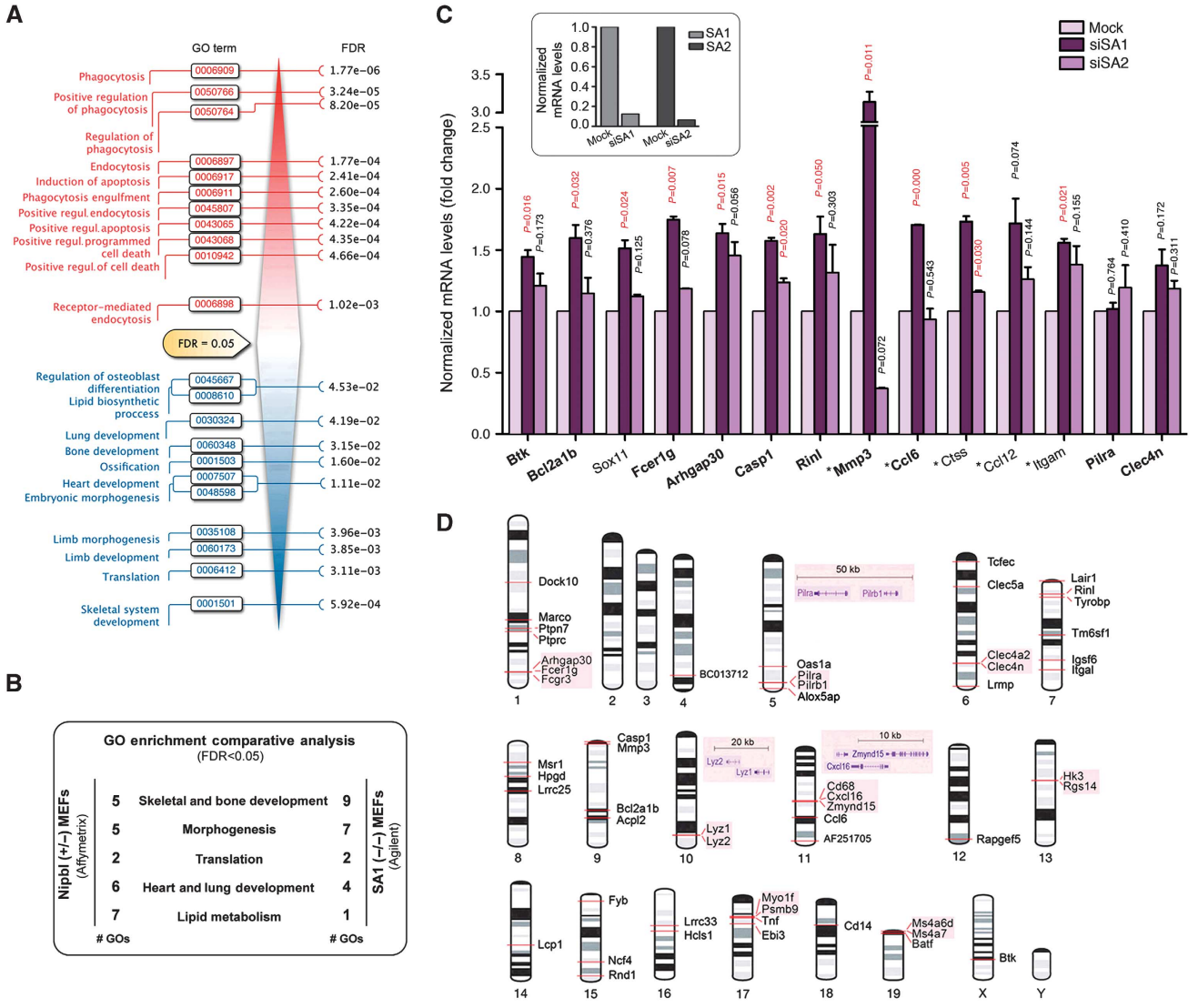


Figure 4 Gene expression changes in the absence of SA1. (**A**) GO analysis (FDR < 0.05) reveals biological processes upregulated (in red) and downregulated (in blue) in SA1-null MEFs compared with wild-type. FDRs for each of the enriched GO terms are indicated. (**B**) GO Comparative Analysis (FDR < 0.05) between transcriptomes from SA1-null and Nipbl-heterozygous MEFs. Common GO terms were grouped in five big biological processes: skeletal and bone development, morphogenesis, translation, heart and lung development and lipid metabolism. The number of GO terms belonging to each group is shown (# GOs). (**C**) Expression of a set of SA1-regulated genes that includes 10 DEGs (shown in bold) and additional genes (Sox11 and skin-related genes, marked with an asterisk) was measured in wild-type MEFs after transfection with SA1 or SA2 siRNAs (results come from triplicate qPCR reactions from two independent experiments). The upper inset shows the efficiency of siRNAs. (**D**) Chromosomal location of the 55 DEGs (FDR < 0.15). Notice that 18 of them are located in close proximity (shadowed).

vance of cohesin-SA1 in positively regulating c-myc. More importantly, ChIP-qPCR performed in brain tissue obtained from wild-type and SA1-null E17.5 embryos indicates that also *in vivo* SA2 does not efficiently replace SA1 (Figure 5B). Myc mRNA and protein levels are reduced in brains from SA1-null embryos, as shown by qPCR and immunostaining (Figure 5C and D, respectively). It is likely that decreased cell proliferation rates due to transcriptional downregulation of c-myc contributes to the lethality of SA1-null embryos. We asked whether other TFs might be regulated by cohesin-SA1. Indeed, gene set analysis revealed that among the transcriptionally altered genes, there is a statistically significant enrichment in genes with binding sites for Pax2 and MafB (Figure 5E, top), two TFs involved in differentiation and development (Cordes and Barsh, 1994; Mansouri *et al*, 1996) whose encoding genes contain cohesin-SA1 at their promoters. Consistently, two of the DEGs are MafB downstream targets (Figure 5E, bottom) and MafB itself is upregulated in SA1-null MEFs (Supplementary Table S3). Therefore, part of the expression changes associated to cohesin-SA1 loss can be secondary to the regulation of genes encoding TFs.

Cohesin-SA1 regulates the expression of gene clusters involved in skin functions

Among DEGs not having cohesin at TSS there are three genes involved in skin development and function (Ccl6, Cxcl16 and Mmp3). Since SA1-null embryos show defects in skin formation (Figure 1B, panel I), we decided to take a closer look at expression changes caused by the absence of SA1 in genes with skin-related functions including immunity, proliferation,

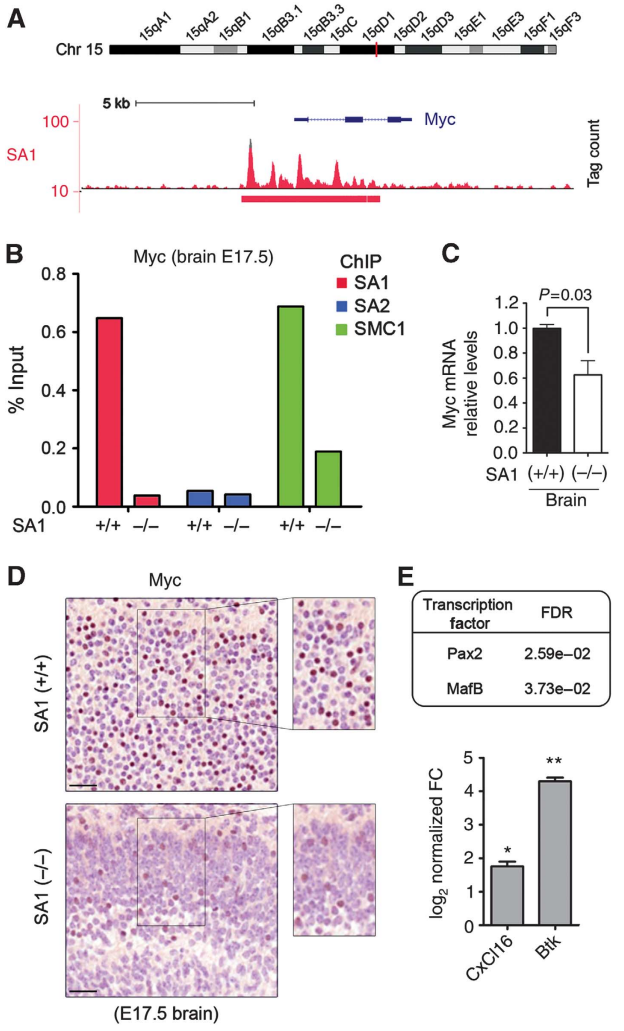


Figure 5 Cohesin-SA1 regulates myc expression. (**A**) SA1-binding region at myc gene (5350 bp) is the widest in the mouse genome (the median is 531 bp). (**B**) Validation by ChIP-qPCR of SA1, SA2 and SMC1 binding at myc promoter in wild-type ($n=12$) and SA1-null ($n=9$) E17.5 brains. (**C**) Myc mRNA levels are significantly reduced in SA1-null brains from E17.5 embryos. Three embryonic brains per genotype were used. (**D**) Immunohistochemistry on E17.5 brains showing reduced Myc protein levels in the cortex of SA1-null embryos. Notice the different cortical structure between wild-type and SA1-null brains. Scale bars, 40 μm. (**E**) Table showing two TFs whose target genes are dysregulated in SA1-null cells. Expression levels of two of those target genes (CxCl16 and Btk) were estimated from three independent qPCR reactions of two clones per genotype. Values are represented as $\log_2$ of fold change (FC) versus wild-type. $^{**}P<0.01$, $^{*}P<0.05$.

skin architecture and motility. Gene Set Enrichment Analysis (GSEA) revealed a highly significant enrichment in those genes in SA1-null MEFs (FDR $<10^{-3}$; Supplementary Figure S6A). Only 15 out of 88 (17%) skin-related genes whose transcription changes have SA1-binding sites up to 5 kb from their TSS (Supplementary Table S7). Strikingly, 63 of the genes (72%) are located in clusters (labelled in blue in Figure 6A), such as the Ccl and Cxcl clusters involved in skin immune function or the keratin clusters involved in skin structure. Analysis of mRNA levels further confirmed the SA1-specific transcriptional changes for a subset of genes implicated in the immune response (Figure 6B; see also genes labelled with an asterisk in Figure 4C). As shown by ChIP-seq for the keratin cluster in chromosome 11 and for the Cxcl cluster in chromosome 5, cohesin distribution within the cluster is considerably altered in SA1-null cells, as judged by the appearance of multiple 'new' sites with low cohesin occupancy (Figure 6C; Supplementary Figure S6B). At the same time, there is much less cohesin at SA1-binding sites, as exemplified by the reduced binding of SMC1 and SMC3 to a site located in the vicinity of Krt15/19 genes in SA1-null cells (indicated with an arrowhead in Figure 6C; qPCR data in Supplementary Figure S2C, upper panels). We therefore hypothesize that cohesin-SA1 plays an architectural role in the organization of these gene clusters that is essential for regulation of their gene expression.

Cohesin-SA1 regulates the expression of Protocadherins in the brain

A detailed analysis of the transcription data revealed the downregulation of many members of the protocadherin (Pcdh) gene family in SA1-null cells. Most, but not all, Pcdh genes are present in three consecutive clusters (Pcdh a, b and g) at mouse chromosome 18. The expression of the Pcdh genes in these clusters is regulated by multiple promoters and alternative *cis* splicing (Yagi, 2008). Our ChIP-seq data identified SA1-binding sites located precisely at most of the multiple TSS of the clustered Pcdh genes (Figure 7A) and also at non-clustered Pcdh genes (e.g., Pcdh7; validation shown in Supplementary Figure S2E). *In vivo*, ChIP-qPCR performed in the brains from wild-type and SA1-null E17.5 embryos showed on the one hand, the prevalence of SA1 over SA2 at the TSS of several of these genes, and on the other, that SA2 does not efficiently replace SA1 in the SA1-null brains, since SMC1 occupancy in these regions is much reduced (Figure 7B). Given that protocadherins are essential for central nervous system development (Morishita and Yagi, 2007), the contribution of cohesin-SA1 to their transcriptional regulation may be particularly critical in the brain. Indeed, we observed a highly significant downregulation of 7 out of 8 Pcdh genes in the brains of E17.5 SA1-null embryos compared with wild-type (Figure 7C). We had chosen these genes because they had also been found downregulated in the brains of Nipbl-heterozygous mice, leading the authors to propose that reduced levels of protocadherins in the brain might contribute to the mental retardation observed in CdLS patients (Kawauchi *et al*, 2009). Importantly, we here show that regulation of Pcdh gene expression in the brain relies on the presence of cohesin-SA1 at their TSS.

Discussion

Vertebrates have two distinct cohesin complexes that contain either SA1 or SA2 (Losada *et al*, 2000; Sumara *et al*, 2000), but most studies assume that both function in a similar way and/or that cohesin-SA2 is more relevant because it is more abundant in somatic cells (Xiao *et al*, 2011). However, ablation of SA1 results in embryonic lethality in mice and late SA1-null mouse embryos present a clear developmental delay (Remeseiro *et al*, 2012; this study). Thus, cohesin-SA1 and cohesin-SA2 are not functionally equivalent. On one hand, we have shown that cohesin-SA1 is responsible for telomere cohesion whereas cohesin-SA2 is critical for centromere cohesion. On the other hand, we report here that cohesin-SA1 has an important role in gene regulation that cannot be assumed by cohesin-SA2.

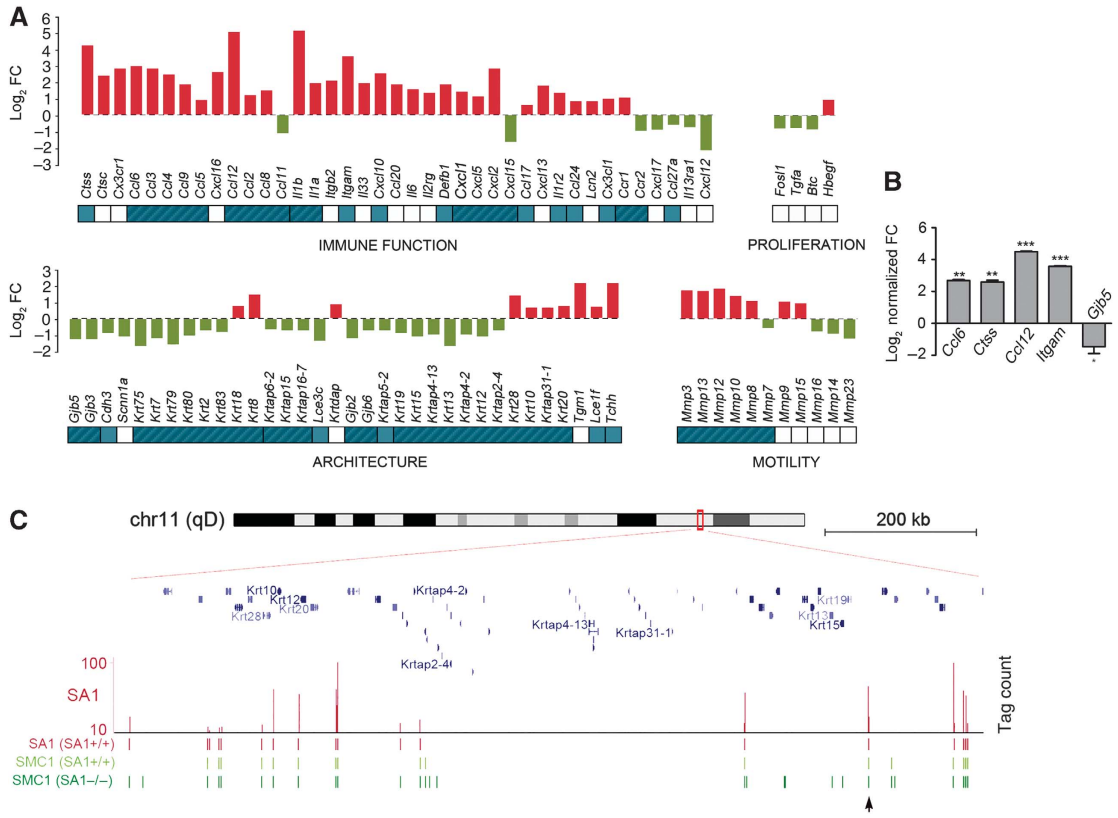


Figure 6 SA1 regulates clustered genes involved in skin development. (**A**) Transcriptional changes detected in genes involved in skin development and function in SA1-null cells. Bars represent the $\log_2$ FC in SA1-null compared with wild-type cells obtained from microarray analysis. Genes in clusters are depicted in blue and genes belonging to the same cluster are grouped (blue shadow). (**B**) Validation by RT–qPCR of transcriptional changes in some of the genes shown in (**A**) (from three independent qPCR reactions of two clones per genotype). $***P<0.001$, $**P<0.01$, $*P<0.05$. (**C**) Detail of SA1 (SA1+/+), SMC1 (SA1+/+) and SMC1 (SA1−/−) binding sites at the Keratin cluster located on chromosome 11. Arrowhead points to a cohesin-SA1-binding site that is validated by ChIP-qPCR in Supplementary Figure S2C.

Differences between SA1 and SA2

We have analysed the genome-wide distributions of SA1 and SA2 in MEFs and found that cohesin-SA1 has a greater propensity for localizing at gene promoters and gene bodies, while cohesin-SA2 prefers intergenic regions. Moreover, in the absence of SA1, cohesin-SA2 (the only cohesin present) redistributes to intergenic regions and fails to accumulate at promoters and other sites nearby genes that could be important for transcriptional regulation, such as sites also bound by CTCF (see model in Figure 8, left). The molecular mechanisms by which cohesin-SA1 regulates gene expression are likely related to the ability of the complex to bring together two DNA sequences not only *in trans* (the sister chromatids), but also *in cis*, thereby facilitating DNA looping (Hadjur *et al*, 2009; Mishiro *et al*, 2009; Nativio *et al*, 2009; Hou *et al*, 2010; Chien *et al*, 2011). These loops may dictate gene expression by promoting or preventing communication between enhancers and promoters in different configurations. One would require the presence of cohesin at the promoter (model in Figure 8, upper right) whereas in other cases, cohesin would be located away from the promoter and would mediate the organization of a gene cluster (Figure 8, lower right). The transcriptional changes observed in SA1-null cells and tissues indicate that cohesin-SA2 cannot fulfil the function of cohesin-SA1. We envision that SA1 and SA2 must confer distinct properties to cohesin. The mechanism by which cohesin associates with chromatin, namely by topological embrace, allows cohesin sliding along DNA (Lengronne *et al*, 2004; Ocampo-Hafalla and Uhlmann, 2011). Cohesin-SA1 could be less prone to sliding than cohesin-SA2 or, in other words, more prone to occupy a fixed genomic position. Cohesion, probably the most important function of cohesin-SA2, does not require localization of cohesin at precise sites, whereas regulation of transcription does. The molecular mechanisms underlying the potential differences in mobility between cohesin-SA1 and cohesin-SA2 remain to be elucidated. One possibility is that the interaction of SA1 and SA2 with Wapl and Pds5, two cohesion factors that modulate the association of cohesin with chromatin throughout the cell cycle, has distinct characteristics (Losada *et al*, 2005; Gandhi *et al*, 2006; Gause *et al*, 2010; Nishiyama *et al*, 2010). Another is the preferential interaction of SA1 with proteins present at defined sites such as CTCF, Mediator or Polycomb (Misulovin *et al*, 2008; Parelho *et al*, 2008; Rubio *et al*, 2008; Wendt *et al*, 2008; Nativio *et al*, 2009; Kagey *et al*, 2010; Strubbe *et al*, 2011; Xiao *et al*, 2011). The identification of proteins that specifically interact with either SA1 or SA2 is one goal of our future work.

Cohesin-SA1 and CdLS

The role of cohesin in transcriptional regulation has become particularly relevant for human health after identification of mutations in the cohesin loader Nipbl and in cohesin subunits as a major cause of CdLS and the subsequent finding that cells from CdLS patients do not show overt cohesion defects (Liu and Krantz, 2009; Dorsett, 2011). A mouse

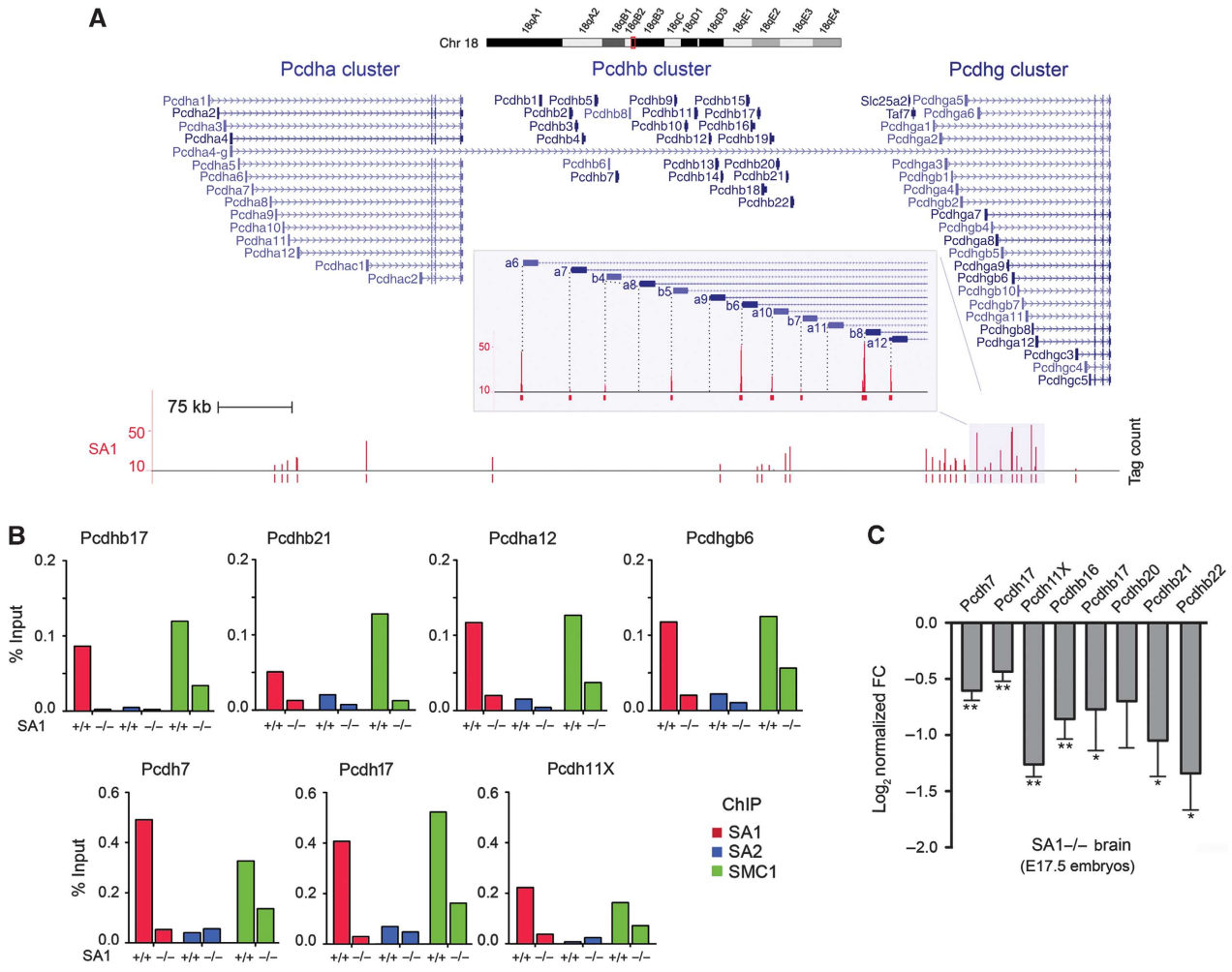


Figure 7 SA1 regulates the expression of protocadherins in mouse brain. (**A**) Detail of SA1-binding sites at Pcdh clusters located in chromosome 18. Notice the position of SA1 at multiple TSS. (**B**) SA1, SA2 and SMC1 binding at the TSS of four clustered and three non-clustered Pcdh genes was validated *in vivo* in wild-type ($n=12$) and SA1-null ($n=9$) E17.5 brains. (**C**) Significant downregulation of Pcdh genes in the brains from E17.5 SA1-null embryos (three embryos per genotype and three independent qPCR reactions per condition). Values are represented as $\log_2$ of FC versus wild-type. $^{**}P<0.01$, $^{*}P<0.05$, Pcdhb20 P-value $=0.13$.

heterozygous for a gene-trap mutation in Nipbl recapitulates some of the pathologies observed in CdLS patients including small size, craniofacial anomalies, heart defects, impaired hearing, delayed bone maturation, reduced body fat and behavioural disturbances (Kawauchi *et al*, 2009). Mouse lacking either Pds5A or Pds5B also exhibit CdLS-like features (Zhang *et al*, 2007, 2009). SA1-heterozygous mice do not show apparent CdLS phenotypes, but SA1-null embryos that survive to late stages of embryogenesis (E17.5–18.5) have clear developmental delay, impaired lipid accumulation and delayed ossification, features that resemble CdLS. Limb or heart defects have not been clearly observed in the limited number of embryos examined. In mice, reduction of Nipbl mRNA levels to 70% is sufficient to elicit CdLS phenotypes without causing apparent cohesion defects (Kawauchi *et al*, 2009). In Drosophila, reduction of Nipped-B expression to 30% of wild-type levels reduces the stable binding of cohesin SA subunit by the same percentage (Gause *et al*, 2010). How the decrease in Nipbl levels affects loading of cohesin-SA1 and cohesin-SA2 on chromatin in mammalian cells remains to be addressed. One scenario is that it affects more acutely the loading of cohesin-SA1 either because it is less abundant (Losada *et al*, 2000; Holzmann *et al*, 2010; Remeseiro *et al*, 2012) or because being more frequently found within genes than cohesin-SA2, cohesin-SA1 may be released from chromatin to allow passage of the transcriptional machinery and thus needs to be constantly reloaded by Nipbl. Alternatively, reduction of Nipbl levels may decrease the amount of both cohesins on chromatin but while this barely affects cohesion, it affects gene transcription (Schaaf *et al*, 2009; Heidinger-Pauli *et al*, 2010). Although we cannot rule out the possibility that cohesin-SA2 affects also gene expression, comparative microarray analysis of human glioblastoma cells expressing or not SA2 shows no evidence for such a role (Solomon *et al*, 2011), consistent with our results in MEFs after downregulation of SA2 by siRNA.

In summary, we have found that cohesin-SA1 plays a unique role in transcriptional regulation that is essential for embryonic development and could underlie the aetiology of CdLS. Although the analysis of SA1-null embryonic fibroblasts combining transcriptional profiling and genome-wide distribution data by ChIP-seq is an important first step, we are aware of the necessity of employing a similar strategy in specific tissues or developmental/differentiation stages in

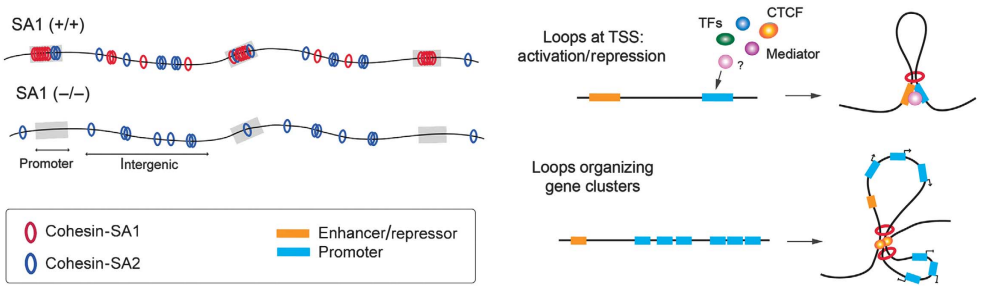


Figure 8 A model for the differential distribution of cohesin-SA1 and cohesin-SA2 and its implications in transcription. Left: Different dynamics and localization of cohesin-SA1 and cohesin-SA2. Cohesin-SA1 is enriched at gene promoters to a much larger extent than cohesin-SA2. In SA1-null cells, cohesin-SA2 fails to accumulate at gene promoters and relocates to intergenic positions. Right: Proposed mechanisms for cohesin-SA1 in regulation of gene expression. Cohesin-SA1 present at gene promoters (upper panel) or in the proximity of genes organized in clusters (lower panel) is required for the formation of loops that arrange the chromatin for gene transcription. The insulator protein CTCF as well as different TFs are likely involved in cohesin-SA1 recruitment to specific genomic positions.

order to understand the contribution of cohesin-SA1 to the establishment of their transcriptomes.

Materials and methods

Identification of SA1, SA2, SMC1 and SMC3 binding sites by ChIP-sequencing

ChIP was performed in SA1-null and wild-type MEFs with custom-made rabbit polyclonal antibodies against SA1, SA2 (described in Remeseiro *et al*, 2012), SMC1 and SMC3 (raised against the C-terminal peptides CEMAKDFVEDDTTHG and CDLTKYPDANPNP-NEQ, respectively), as described (Cuadrado *et al*, 2010) with some modifications. Cells were cross-linked by addition of 1% formaldehyde for 10 min at room temperature, harvested and frozen in liquid nitrogen. Cells were then resuspended in lysis buffer (2×10^7 cells/ml) and sonicated on Covaris system (shearing time 30 min, 20% duty cycle, intensity 10, 200 cycles per burst and 30 s per cycle) in a volume of 2 ml. From 4 to 15 ng of immunoprecipitated chromatin (as quantitated by fluorometry) were electrophoresed on agarose gel and independent sample-specific fractions of 100–200 bp were taken. These samples were processed into sequencing libraries and analysed according to Illumina's 'ChIP-Sequencing Sample Prep Guide' (part #11257047 Rev. A), with the exception that gel extraction was replaced with Agencourt AMPure XP (Beckman Coulter) bead purification. Adapter-ligated library was completed by limited-cycle PCR with Illumina PE primers (14 cycles). DNA libraries were applied to an Illumina flow cell for cluster generation and sequenced on the Illumina Genome Analyzer IIx (GAIIx). Image analysis was performed with Illumina Real Time Analysis software (RTA1.8). Sequence alignment to the reference genome (NCBIm37/mm9, April 2007) was made with Illumina's ELANDv2 algorithm on its 'eland_extended' mode from within CASAVA-1.7 package, using default settings. Only reads with a unique alignment in the reference genome were used for the peak detection, which was performed using MACS v1.4 setting a FDR < 0.1 and a *P*-value cutoff of 10^{-5} (Zhang *et al*, 2008). All comparisons were done using the input tracks as control, and each one of the data sets as treatment, using the following naming convention: SA1 (SA1 + / +), SA2 (SA1 + / +), SMC1 (SA1 + / +), SMC3 (SA1 + / +), SA1 (SA1−/−), SA2 (SA1−/−), SMC1 (SA1−/−) and SMC3 (SA1−/−). In the case of SA1 (SA1 + / +) data set, the 176 peaks obtained from SA1 ChIP in SA1−/− cells were cleared. CTCF peaks were obtained from the UCSC Genome Browser track 'MEL CTCF D Pk', part of the ENCODE/Stanford/Yale data set of TF-binding sites (TFBS) in mouse. Genomic interval overlaps and signal distributions were obtained using BEDTools v2.12 (Quinlan and Hall, 2010), PeakAnalyzer v1.3 (Salmon-Divon *et al*, 2010) and custom UNIX shell scripting. All the statistical tests and correlations were calculated using R functions (http://cran.r-project.org).

ChIP-qPCR in embryonic brain

The brains from E17.5 embryos were extracted and minced in cold PBS with protease inhibitors cocktail (Roche #11873580001). Small tissue pieces (< 1 mm) were then cross-linked for 20 min RT in fixing solution (1% formaldehyde, 50 mM HEPES-KOH, 100 mM NaCl, 1 mM EDTA and 0.5 mM EGTA). Cross-linking was stopped by adding 1/20 volume of 2.5 M Glycine for 5 min at RT. After two washes in PBS, tissues were resuspended in lysis buffer and processed for further ChIP analysis as described in the above section.

Validation of ChIP-seq results

ChIP-seq results were validated by ChIP-qPCR on immunoprecipitated chromatin from two clones per genotype. Chromosome coordinates of the validated peaks and the corresponding primers are listed in Supplementary Table S8. The relative amount of each amplified fragment was normalized with respect to the amplification obtained from input DNA and represented as percentage of 1 μg of input DNA.

Gene expression analysis

Total RNA from three wild-type and three SA1-null MEF clones was analysed by two-colour hybridization on Whole Mouse Genome DNA microarrays (G4122F; Agilent), and images were quantified with Agilent Feature Extraction Software (v. 10.1.1). DEGs between SA1-null and wild-type MEFs were obtained by limma (Smyth GK; Bioconductor project; http://www.bioconductor.org). FDR adjustment was employed to account for multiple testing. Raw data from gene expression microarray experiments in Nipbl + /− MEFs were kindly provided by the authors (Kawauchi *et al*, 2009). In this case, Affymetrix Murine 430A 2.0 were normalized using Robust Multi-array Average (RMA) algorithm available in Bioconductor's affy package and limma package was used to obtain DEGs.

Functional analysis for GO terms

GO analyses at Biological Process, Cellular Component and Molecular Function were performed using FatiScan tool available at Babelomics suite (http://www.babelomics.org). To this end, genes were ranked based on limma's moderated *t* statistic and GO enrichment was evaluated by segmentation test. GO terms showing FDR < 0.05 were considered statistically significant.

Gene set analysis of Myc targets and skin-related genes

GSEA (Subramanian *et al*, 2005) was employed to evaluate the enrichment of custom gene sets in our microarrays experiments. Myc targets were obtained from literature (Chen *et al*, 2008; Kim *et al*, 2008, 2010; Sridharan *et al*, 2009; Smith *et al*, 2010) whereas skin gene set was built from Nagarajan *et al* (2010) and references therein. GSEA was run using gene expression values ranked by limma moderated *t* statistic. After Kolmogorov–Smirnoff testing, those gene sets showing FDR < 0.1, were considered enriched between SA1-null and wild-type MEFs.

Enrichment analysis for target genes of TFs

GSEA for Jaspar TFBS was done using Fatiscan tool available at Babelomics platform (http://www.babelomics.org). Genes were

ranked by limma moderated *t* statistic. TFBS showing FDR <0.05 were considered enriched between SA1-null and wild-type MEFs.

mRNA isolation and quantitative real-time PCR (qRT–PCR) analysis

Total RNA was isolated from MEFs using RNeasy Kit (Qiagen) and cDNA was synthesized with SuperScript™ II reverse transcriptase (Invitrogen) using random hexamer primers. An Applied Biosystems 7900HT Fast qRT–PCR was used to determine mRNA levels. GAPDH was used for normalization. Primers used for mRNA amplification are described in Supplementary Table S9.

RNA interference, immunoprecipitation and immunoblotting

Interference of SA1 and SA2 was performed with siGENOME SMARTpool siRNAs from Dharmacon (M-041989 and M-057033, respectively) at a final concentration of 100 nM and using DharmaFECT transfection reagent 1 in the case of C2C12 cells and the Neon transfection system (Invitrogen) in the case of MEFs. ChIP with SA1- and SA2-specific antibodies was performed 72–96 h after transfection. Immunoprecipitation was carried out with Nuclear Complex Co-IP Kit (Active Motif, 54001) from cell extracts according to the manufacturer's instructions, with SA1, SA2, SMC1 and SMC3 specific antibodies. Whole-cell extracts were prepared by lysing and sonicating a cell pellet in SDS–PAGE loading buffer and equal amounts of protein were run in 7.5% Bis/Tris gels followed by western blotting.

Histology and immunohistochemistry

E17.5 embryos were fixed in 10% buffered formalin (Sigma) and embedded in paraffin using standard procedures. In all, 3 μm sections were stained with haematoxylin and eosin (HE) and subjected to histopathological analysis. Anti-myc (Santa Cruz, sc-764), anti-SA1 and anti-SA2 were used for immunohistochemical analysis of 3 μm sections. Positive cells were visualized using 3,3-diaminobenzidine tetrahydrochloride plus (DAB+) as a chromogen, and counterstaining was performed with haematoxylin.

Statistical analysis

GraphPad Prism 5 software was used to calculate two-tailed χ^2 test (with Yates' correction) in Figure 3B, Mann–Whitney *U*-test in Figure 3E and two-tailed Student's *t*-test in Figures 4C, 5C, 5E, 6B and 7C, Supplementary Figures S1C and S4. Limma package was used for GO, GSEA and gene expression analysis to obtain DEGs.

Data access

Microarray and ChIP-sequencing data from this study have been submitted to GEO database and have been approved with the following Series reference: GSE32320 (GSE32234 and GSE32319 SubSeries correspond to microarray and ChIP-seq data, respectively).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We are very grateful to M Cañamero (Pathology Unit, CNIO) for histopathological analyses and to O Domínguez (Genomics Unit, CNIO) for help and advise on ChIP-seq. We also thank M Serrano and O Fernández-Capetillo for critically reading the manuscript. Research in our laboratory is supported by the Spanish Ministry of Science and Innovation (SAF-2010-21517 and CSD2007-00015 to AL; Ramón y Cajal grant for AC). SR is the recipient of a 'La Caixa' predoctoral fellowship.

Author contributions: AL designed and supervised the study. SR and AC designed, performed, analysed and interpreted all the experiments. DGP and GGL performed and interpreted the ChIP-seq and microarray data analysis, respectively. AL, SR and AC wrote the manuscript with ideas and comments from GGL and DGP.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Bose T, Gerton JL (2010) Cohesinopathies, gene expression, and chromatin organization. *J Cell Biol* **189:** 201–210

Canudas S, Smith S (2009) Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells. *J Cell Biol* **187:** 165–173

Castronovo P, Gervasini C, Cereda A, Masciadri M, Milani D, Russo S, Selicorni A, Larizza L (2009) Premature chromatid separation is not a useful diagnostic marker for Cornelia de Lange syndrome. *Chromosome Res* **17:** 763–771

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK *et al* (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133:** 1106–1117

Chien R, Zeng W, Kawauchi S, Bender MA, Santos R, Gregson HC, Schmiesing JA, Newkirk DA, Kong X, Ball Jr AR, Calof AL, Lander AD, Groudine MT, Yokomori K (2011) Cohesin mediates chromatin interactions that regulate mammalian beta-globin expression. *J Biol Chem* **286:** 17870–17878

Cordes SP, Barsh GS (1994) The mouse segmentation gene kr encodes a novel basic domain-leucine zipper transcription factor. *Cell* **79:** 1025–1034

Cuadrado A, Corrado N, Perdiguero E, Lafarga V, Munoz-Canoves P, Nebreda AR (2010) Essential role of p18Hamlet/SRCAP-mediated histone H2A.Z chromatin incorporation in muscle differentiation. *EMBO J* **29:** 2014–2025

Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodriguez C, Arnedo M, Loeys B, Kline AD, Wilson M, Lillquist K, Siu V, Ramos FJ, Musio A, Jackson LS, Dorsett D, Krantz ID (2007) Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. *Am J Hum Genet* **80:** 485–494

Degner SC, Wong TP, Jankevicius G, Feeney AJ (2009) Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J Immunol* **182:** 44–48

Dorsett D (2011) Cohesin: genomic insights into controlling gene transcription and development. *Curr Opin Genet Dev* **21:** 199–206

Dorsett D, Eissenberg JC, Misulovin Z, Martens A, Redding B, McKim K (2005) Effects of sister chromatid cohesion proteins on cut gene expression during wing development in Drosophila. *Development* **132:** 4743–4753

Gandhi R, Gillespie PJ, Hirano T (2006) Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Curr Biol* **16:** 2406–2417

Gause M, Misulovin Z, Bilyeu A, Dorsett D (2010) Dosage-sensitive regulation of cohesin chromosome binding and dynamics by Nipped-B, Pds5, and Wapl. *Mol Cell Biol* **30:** 4940–4951

Guillou E, Ibarra A, Coulon V, Casado-Vela J, Rico D, Casal I, Schwob E, Losada A, Mendez J (2010) Cohesin organizes chromatin loops at DNA replication factories. *Genes Dev* **24:** 2812–2822

Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M (2009) Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* **460:** 410–413

Hallson G, Syrzycka M, Beck SA, Kennison JA, Dorsett D, Page SL, Hunter SM, Keall R, Warren WD, Brock HW, Sinclair DA, Honda BM (2008) The Drosophila cohesin subunit Rad21 is a trithorax group (trxG) protein. *Proc Natl Acad Sci USA* **105:** 12405–12410

Heidinger-Pauli JM, Mert O, Davenport C, Guacci V, Koshland D (2010) Systematic reduction of cohesin differentially affects chromosome segregation, condensation, and DNA repair. *Curr Biol* **20:** 957–963

Holzmann J, Fuchs J, Pichler P, Peters JM, Mechtler K (2010) Lesson from the stoichiometry determination of the cohesin complex: a

short protease mediated elution increases the recovery from cross-linked antibody-conjugated beads. *J Proteome Res* **10:** 780–789

Horsfield JA, Anagnostou SH, Hu JK, Cho KH, Geisler R, Lieschke G, Crosier KE, Crosier PS (2007) Cohesin-dependent regulation of Runx genes. *Development* **134:** 2639–2649

Hou C, Dale R, Dean A (2010) Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc Natl Acad Sci USA* **107:** 3651–3656

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467:** 430–435

Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, Nussenzweig A, Kitzes L, Yokomori K, Hallgrimsson B, Lander AD (2009) Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/−) mouse, a model of Cornelia de Lange Syndrome. *PLoS Genet* **5:** e1000650

Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132:** 1049–1061

Kim J, Woo AJ, Chu J, Snow JW, Fujiwara Y, Kim CG, Cantor AB, Orkin SH (2010) A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* **143:** 313–324

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H, Bamshad MJ, Carey JC, Rappaport E, Kawauchi S, Lander AD, Calof AL, Li HH, Devoto M *et al* (2004) Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. *Nat Genet* **36:** 631–635

Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, Itoh T, Watanabe Y, Shirahige K, Uhlmann F (2004) Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430:** 573–578

Lin W, Jin H, Liu X, Hampton K, Yu HG (2011) Scc2 regulates gene expression by recruiting cohesin to the chromosome as a transcriptional activator during yeast meiosis. *Mol Biol Cell* **22:** 1985–1996

Liu J, Krantz ID (2009) Cornelia de Lange syndrome, cohesin, and beyond. *Clin Genet* **76:** 303–314

Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, Spinner NB, Vega H, Jackson LG, Shirahige K, Krantz ID (2009) Transcriptional dysregulation in NIPBL and cohesin mutant human cells. *PLoS Biol* **7:** e1000119

Losada A, Hirano T (2005) Dynamic molecular linkers of the genome: the first decade of SMC proteins. *Genes Dev* **19:** 1269–1287

Losada A, Yokochi T, Hirano T (2005) Functional contribution of Pds5 to cohesin-mediated cohesion in human cells and Xenopus egg extracts. *J Cell Sci* **118:** 2133–2141

Losada A, Yokochi T, Kobayashi R, Hirano T (2000) Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. *J Cell Biol* **150:** 405–416

Mansouri A, Hallonet M, Gruss P (1996) Pax genes and their roles in cell differentiation and development. *Curr Opin Cell Biol* **8:** 851–857

Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M (2009) Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO J* **28:** 1234–1245

Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, Macarthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, Dorsett D (2008) Association of cohesin and Nipped-B with transcriptionally active regions of the Drosophila melanogaster genome. *Chromosoma* **117:** 89–102

Morishita H, Yagi T (2007) Protocadherin family: diversity, structure, and function. *Curr Opin Cell Biol* **19:** 584–592

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L (2006) X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. *Nat Genet* **38:** 528–530

Muto A, Calof AL, Lander AD, Schilling TF (2011) Multifactorial origins of heart and gut defects in nipbl-deficient zebrafish, a model of Cornelia de Lange Syndrome. *PLoS Biol* **9:** e1001181

Nagarajan P, Chin SS, Wang D, Liu S, Sinha S, Garrett-Sinha LA (2010) Ets1 blocks terminal differentiation of keratinocytes and induces expression of matrix metalloproteases and innate immune mediators. *J Cell Sci* **123:** 3566–3575

Nasmyth K, Haering CH (2009) Cohesin: its roles and mechanisms. *Annu Rev Genet* **43:** 525–558

Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters JM, Murrell A (2009) Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet* **5:** e1000739

Nishiyama T, Ladurner R, Schmitz J, Kreidl E, Schleiffer A, Bhaskara V, Bando M, Shirahige K, Hyman AA, Mechtler K, Peters JM (2010) Sororin mediates sister chromatid cohesion by antagonizing Wapl. *Cell* **143:** 737–749

Ocampo-Hafalla MT, Uhlmann F (2011) Cohesin loading and sliding. *J Cell Sci* **124:** 685–691

Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132:** 422–433

Pauli A, Althoff F, Oliveira RA, Heidmann S, Schuldiner O, Lehner CF, Dickson BJ, Nasmyth K (2008) Cell-type-specific TEV protease cleavage reveals cohesin functions in Drosophila neurons. *Dev Cell* **14:** 239–251

Pauli A, van Bemmel JG, Oliveira RA, Itoh T, Shirahige K, van Steensel B, Nasmyth K (2010) A direct role for cohesin in gene regulation and ecdysone response in Drosophila salivary glands. *Curr Biol* **20:** 1787–1798

Peters JM, Tedeschi A, Schmitz J (2008) The cohesin complex and its roles in chromosome biology. *Genes Dev* **22:** 3089–3114

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842

Remeseiro S, Cuadrado A, Carretero M, Martínez P, Drosopoulos WC, Cañamero M, Schildkraut CL, Blasco MA, Losada A (2012) Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres. *EMBO J* **31:** 2076–2089

Rhodes JM, Bentley FK, Print CG, Dorsett D, Misulovin Z, Dickinson EJ, Crosier KE, Crosier PS, Horsfield JA (2010) Positive regulation of c-Myc by cohesin is direct, and evolutionarily conserved. *Dev Biol* **344:** 637–649

Rollins RA, Korom M, Aulner N, Martens A, Dorsett D (2004) Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. *Mol Cell Biol* **24:** 3100–3111

Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A (2008) CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci USA* **105:** 8309–8314

Salmon-Divon M, Dvinge H, Tammoja K, Bertone P (2010) PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics* **11:** 415

Schaaf CA, Misulovin Z, Sahota G, Siddiqui AM, Schwartz YB, Kahn TG, Pirrotta V, Gause M, Dorsett D (2009) Regulation of the Drosophila enhancer of split and invected-engrailed gene complexes by sister chromatid cohesion proteins. *PLoS One* **4:** e6202

Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT (2010) A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res* **20:** 578–588

Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, Marks H, Adams DJ, Schatz DG, Aragon L, Fisher AG, Krangel MS, Nasmyth K, Merkenschlager M (2011) A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* **476:** 467–471

Smith KN, Singh AM, Dalton S (2010) Myc represses primitive endoderm differentiation in pluripotent stem cells. *Cell Stem Cell* **7:** 343–354

Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, Harris BT, Toretsky JA, Rosenberg SA, Shukla N, Ladanyi M, Samuels Y, James CD, Yu H, Kim JS, Waldman T (2011) Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science* **333:** 1039–1043

Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S, Zhou Q, Plath K (2009) Role of the murine reprogramming factors in the induction of pluripotency. *Cell* **136:** 364–377

Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM (2008) Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J* **27:** 654–666

Strubbe G, Popp C, Schmidt A, Pauli A, Ringrose L, Beisel C, Paro R (2011) Polycomb purification by *in vivo* biotinylation tagging reveals cohesin and Trithorax group proteins as interaction partners. *Proc Natl Acad Sci USA* **108:** 5572–5577

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102:** 15545–15550

Sumara I, Vorlaufer E, Gieffers C, Peters BH, Peters JM (2000) Characterization of vertebrate cohesin complexes and their regulation in prophase. *J Cell Biol* **151:** 749–762

Tonkin ET, Smith M, Eichhorn P, Jones S, Imamwerdi B, Lindsay S, Jackson M, Wang TJ, Ireland M, Burn J, Krantz ID, Carr P, Strachan T (2004) A giant novel gene undergoing extensive alternative splicing is severed by a Cornelia de Lange-associated translocation breakpoint at 3q26.3. *Hum Genet* **115:** 139–148

Trumpp A, Refaeli Y, Oskarsson T, Gasser S, Murphy M, Martin GR, Bishop JM (2001) c-Myc regulates mammalian body size by controlling cell number but not cell size. *Nature* **414:** 768–773

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451:** 796–801

Xiao T, Wallace J, Felsenfeld G (2011) Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol Cell Biol* **31:** 2174–2183

Yagi T (2008) Clustered protocadherin family. *Dev Growth Differ* **50**(Suppl 1)**:** S131–S140

Zhang B, Chang J, Fu M, Huang J, Kashyap R, Salavaggione E, Jain S, Kulkarni S, Deardorff MA, Uzielli ML, Dorsett D, Beebe DC, Jay PY, Heuckeroth RO, Krantz I, Milbrandt J (2009) Dosage effects of cohesin regulatory factor PDS5 on mammalian development: implications for cohesinopathies. *PLoS One* **4:** e5232

Zhang B, Jain S, Song H, Fu M, Heuckeroth RO, Erlich JM, Jay PY, Milbrandt J (2007) Mice lacking sister chromatid cohesion protein PDS5B exhibit developmental abnormalities reminiscent of Cornelia de Lange syndrome. *Development* **134:** 3191–3201

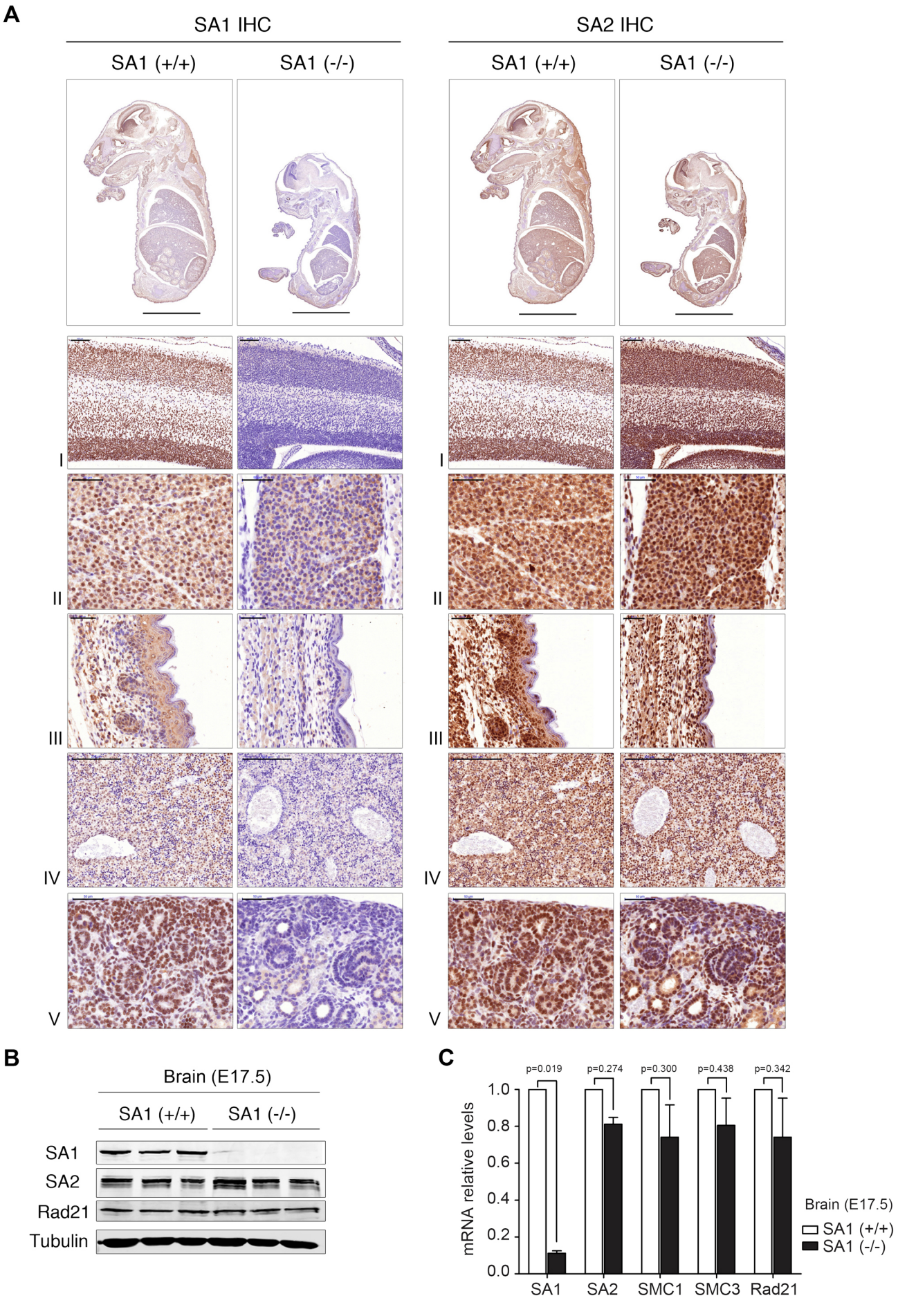
Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.

Figure S1. SA2 expression remains unchanged in SA1-null embryos.

(A) Immunohistochemistry with SA1 and SA2 antibodies performed on tissue sections from wild-type and SA1-null E17.5 embryos. Scale bars, 5mm. I. Brain cortex. Scale bars, 100 μm. II. Interscapular adipose tissue. Scale bars, 50 μm. III. Skin. Scale bars, 50 μm. IV. Liver. Scale bars, 200 μm. V. Kidney. Scale bars, 50 μm. **(B)** Immunoblot analysis of brains from wild-type and SA1-null E17.5 embryos with the indicated antibodies. Tubulin was used as loading control. **(C)** mRNA levels of all cohesin complex subunits in brains from SA1-null E17.5 embryos relative to the wild-type littermates (3 embryos per genotype and 3 independent qPCR reactions per condition).

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.

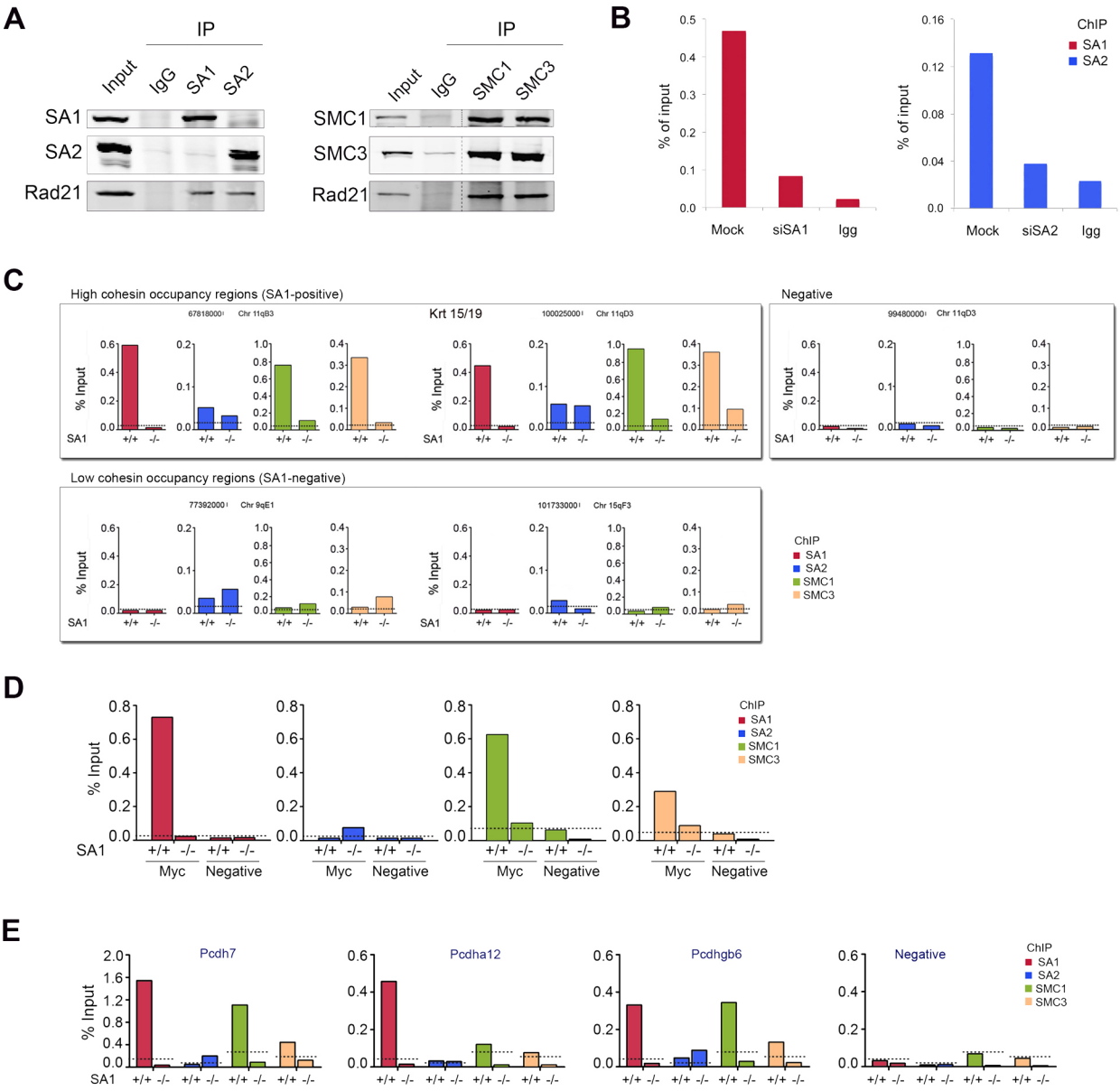


Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.

Figure S2. Chip-seq validation.

(A) Antibodies against SA1, SA2, SMC1 and SMC3 used for ChIP were previously tested for immunoprecipitation using a mouse cell extract enriched in chromatin-bound proteins. Immunoblot analysis with an antibody against Rad21, a subunit common to both cohesin-SA1 and cohesin-SA2, was used to assess the efficiency of the immunoprecipitation reactions. **(B)** Antibody validation for ChIP was done by ChIP-qPCR analysis of SA1 and SA2 binding to mouse IFNg locus (Hadjur et al., 2009), performed with chromatin from C2C12 mouse myoblasts treated with either mock, SA1 or SA2 siRNAs. **(C)** ChIP-qPCR validation of binding sites identified for SA1, SA2, SMC1 and SMC3 by ChIP-sequencing in wild-type and SA1-null MEFs . Regions with high cohesin occupancy (SA1 positive), low cohesin occupancy (SA1 negative), and negative for cohesin (negative) are shown. Dotted lines indicate the background level (cohesin binding in a negative region) for each antibody. Same analysis was performed for myc promoter and an adjacent negative region **(D)** and for Pcdh genes **(E).** Bars in (C-E) represent the average of two independent experiments performed with two different clones of MEFs per genotype.

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.

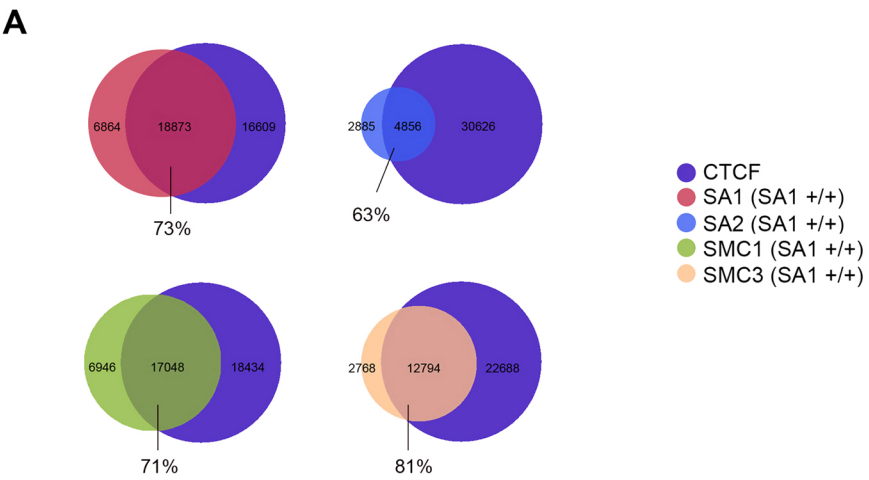


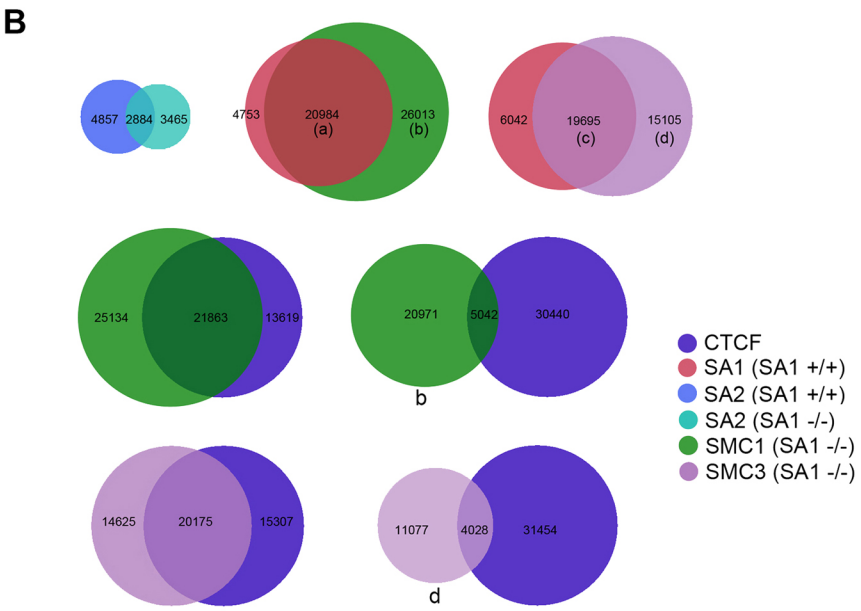
Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.

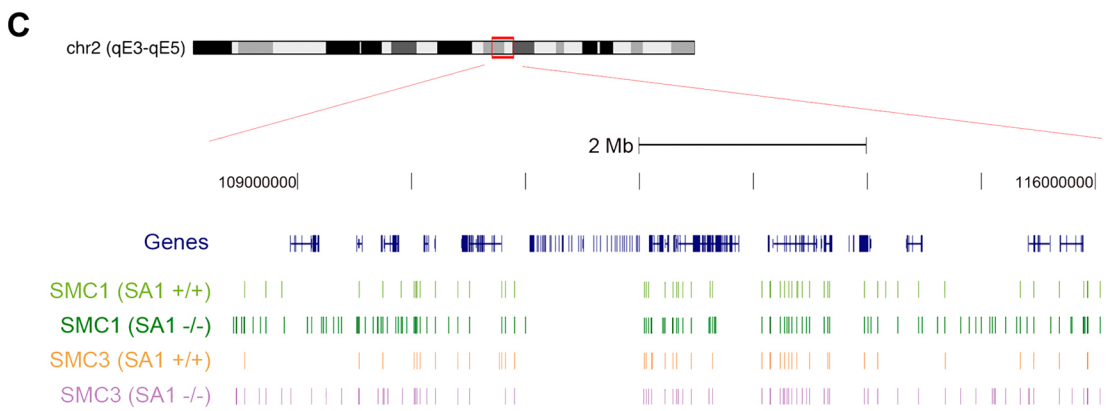
Figure S3. Peak overlapping analysis.

(A) Venn diagrams showing the overlap between the binding sites identified for each cohesin subunit (SA1, SA2, SMC1 and SMC3) and CTCF (from Yale Dataset) in wild-type cells. **(B)** Additional Venn diagrams showing the overlap between the indicated groups of binding sites. **(C)** Representative image of the browser showing the redistribution of SMC1 and SMC3 in SA1-null cells with respect to wild-type cells.

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.







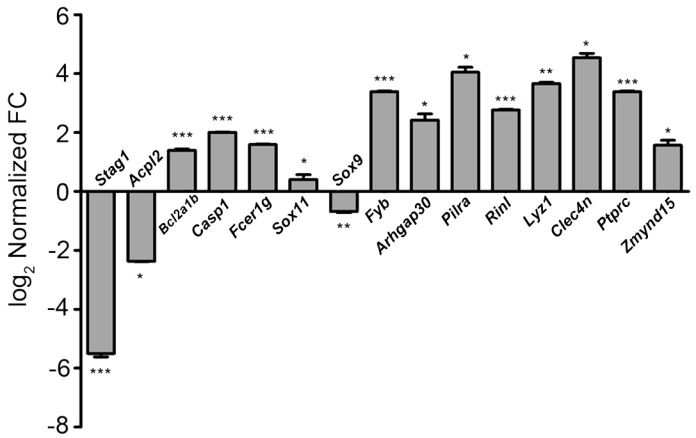
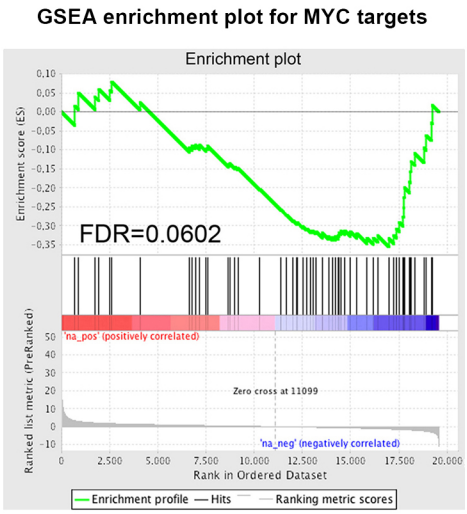


Figure S4. Validation of microarray data.

Validation of microarray data from wild-type and SA1-null MEFs performed by RT-qPCR (2 different clones per genotype and 3 independent qPCR reactions). Values are represented as $\log_2$ of fold change (FC) versus wild-type. For statistical analysis, a two-tailed Student's t-test was done (***p<0.001, **p<0.01, *p<0.05).

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.



MYC targets gene set

Prdx1	*Pla2g6*	*Abce1*
Tardbp	*Rplp1*	*Mat2b*
Incenp	*Naca*	*Tcfap4*
Rps27l	*Mcm5*	*Iars*
Atpif1	*Bc031781*	*Nolc1*
Rpl7	*Rps19*	*Rsl1d1*
Sp1	*Uhrf1*	*Nvl*
Slc9a8	*Gars*	*Mthfd1*
Laptm4a	*Arv1*	*Rps25*
Mrpl50	*Sf3a1*	*Nmd3*
Zcchc7	*Cct7*	*Ddb1*
Bre	*Kif20a*	*Luc7l*
Rps3	*Rplp2*	*Nek8*
Slc3a2	*Ubap2*	*Tcp1*
Snrpa	*Atad3a*	*Eef2*
St13	*Xpo5*	*Ppan*
Hspd1	*Eef1b2*	*Ddx39*
Nudcd2	*Ranbp1*	*Rpl32*
Cstf1	*Cct4*	*Atf4*
Ebna1bp2	*Paics*	

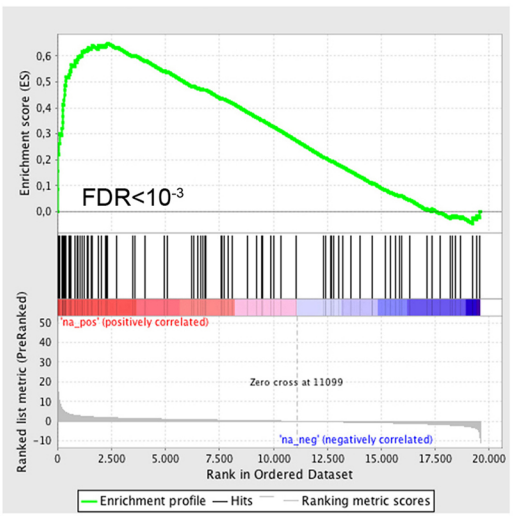
Figure S5. GSEA analysis of Myc target genes.

Gene set enrichment plot of the indicated Myc targets tested in microarrays from SA1-null and wild-type MEFs.

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.

A

GSEA enrichment plots for skin gene set



Skin gene set

Areg	*Cxcl13*	*Il18rap*	*Krt25*	*Lce1b*	*S100a5*
Btc	*Cxcl14*	*Il18rb*	*Krt27*	*Lce1c*	*S100a8*
Ccl11	*Cxcl16*	*Il1a*	*Krt28*	*Lce1d*	*S100b*
Ccl12	*Cxcl2*	*Il1b*	*Krt32*	*Lce1f*	*Scnn1a*
Ccl19	*Cxcl5*	*Il1f8*	*Krt6a*	*Lce1h*	*Sprr1a*
Ccl2	*Cxcl9*	*Il1r2*	*Krt6b*	*Lce1i*	*Sprr1b*
Ccl20	*Defb1*	*Il2ra*	*Krt7*	*Lce3a*	*Sprr2a*
Ccl3	*Defb3*	*Il2rg*	*Krt71*	*Lce3c*	*Sprr2d*
Ccl4	*Defb6*	*Il33*	*Krt75*	*Lcn2*	*Sprr2f*
Ccl5	*Epgn*	*Il4ra*	*Krt77*	*Lcn8*	*Sprr2i*
Ccl6	*Fil*	*Il6*	*Krt79*	*Lor*	*Sprr2j*
Ccl8	*Flg*	*Itgal*	*Krt8*	*Lpo*	*Sprr2k*
Ccl9	*Fos*	*Itgam*	*Krt80*	*Mmp10*	*Tchh*
Ccr1	*FosL1*	*Itgb2*	*Krt83*	*Mmp12*	*Tgfa*
Ccr2	*FosL2*	*JunB*	*Krtap15*	*Mmp13*	*Tgfb1*
Cdh3	*Gjb2*	*Krt10*	*Krtap16-7*	*Mmp15*	*Tgm1*
Cldn23	*Gjb3*	*Krt12*	*Krtap2-4*	*Mmp1a*	*Tgm3*
Cldnd1	*Gjb4*	*Krt13*	*Krtap31-1*	*Mmp1b*	*Timb1*
Cts8	*Gjb5*	*Krt15*	*Krtap4-13*	*Mmp3*	*Tslp*
Ctsc	*Gjb6*	*Krt16*	*Krtap4-2*	*Mmp7*	
Ctss	*Hbegf*	*Krt17*	*Krtap5-2*	*Mmp8*	
Cx3cl1	*Il133ra1*	*Krt18*	*Krtap6-2*	*Mmp9*	
Cx3cr1	*Il17a*	*Krt19*	*Krtdap*	*Prss8*	
Cxcl10	*Il17d*	*Krt2*	*Lce1a1*	*S100a1*	
Cxcl11	*Il18*	*Krt20*	*Lce1a2*	*S100a4*	

B

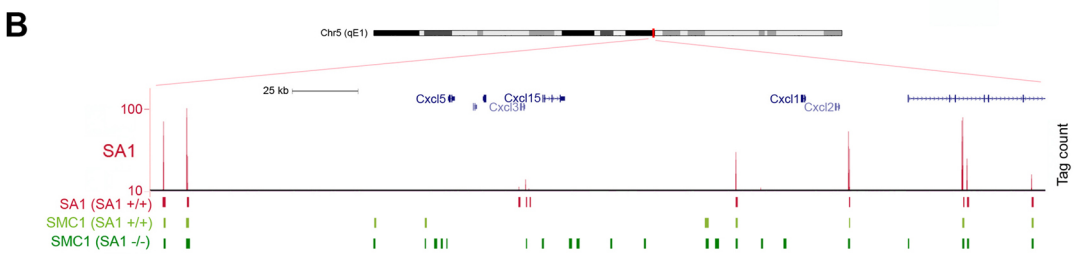


Figure S6. Cohesin-SA1 regulates skin-related genes.

(A) Gene set enrichment plot of the indicated skin-related genes tested in microarrays from SA1-null and wild-type MEFs. **(B)** Distribution of cohesin binding sites at the Cxcl cluster located on chromosome 5.

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.

Supplementary Table S1: ChIP-seq experiment summary. Overall reads obtained in Illumina ChIP-seq experiments. Alignments on NCBIm37 genome assembly were made with Illumina's Eland software (GAPileline 1.4.0), allowing no more than 2 mismatches within 36 base seed lenghts.

Sample	Sequenced reads (QC PASS)	Uniquely aligned reads	% uniquely aligned	Duplicate Filtered Tags (Redundant)
Input 1	30,837,796	21,931,673	71	21,298,545 (3%)
SA1 (SA1 +/+)	33,045,532	22,412,267	68	21,383,266 (5%)
SA2 (SA1 +/+)	34,511,610	21,521,101	62	19,456,228 (10%)
SA2 (SA1 -/-)	35,015,268	22,899,617	65	21,850,166 (5%)
Input 2	38,619369	26,929,896	70	26,239,283 (3%)
SA1 (SA1 -/-)	35,702,433	20,126,016	56	19,465,428 (3%)
SMC1 (SA1 +/+)	34,490,522	21,948,219	64	21,375,089 (3%)
SMC1 (SA1 -/-)	39,771,621	21,088,100	53	20,147,423 (4%)
SMC3 (SA1 +/+)	38,659,833	24,690,999	64	23,975,691 (3%)
SMC3 (SA1 -/-)	37,835,948	26,563,680	70	25,820,540 (3%)

Silvia Remeseiro*, Ana Cuadrado*, Gonzalo Gómez-López, David G Pisano and Ana Losada. **A unique role of cohesin-SA1 in gene regulation and development.** The EMBO Journal (2012) 31, 2090–2102.

Supplementary Table S2: ChIP-seq experiment, cleaning criteria. Peaks from chromosome Y and Mitochondrial DNA have been discarded. Inputs have been used as control condition in all experiments. Any region enriched in inputs has been eliminated. SA1 signals from SA1-/- have been removed from SA1+/+.

Sample	Regions/Peaks (FDR≤10%)	Input cleaned	SA1-/- cleaned	Double Enrichment
Input 1	436	0		
SA1 (SA1 +/+)	26,602	26,600	26,594	25,737
SA2 (SA1 +/+)	7,741	7,741		
SA2 (SA1 -/-)	6,352	6,349		
Input 2	417	0		
SA1 (SA1 -/-)	177	176		
SMC1 (SA1 +/+)	24,000	23,994		
SMC1 (SA1 -/-)	47,003	46,997		
SMC3 (SA1 +/+)	15,564	15,562		
SMC3 (SA1 -/-)	34,802	34,800		

Sample	Regions/Peaks (FDR≤10%)
CTCF_MEL_YALE	35,482

***Due to space constrains, all the supplementary tables not shown here can be found in the CD copy.**

The specific contributions of cohesin-SA1 to cohesion and gene expression: Implications for cancer and development

Ana Cuadrado, **Silvia Remeseiro,** Gonzalo Gómez-López, David G Pisano and Ana Losada.

Cell Cycle (2012) 11:12, 2233-2238

In this paper we summarize the major finding of the two previous publications and further explore the implication of cohesin-SA1 in CdLS. We had previously shown that Gene Ontology (GO)-defined processes related to abnormalities found in CdLS patients were downregulated in SA1-null cells, consistently with the observation of CdLS features in late SA1-null embryos. A highly significant overlap was also observed when we compared these altered processes and those from the Nipbl heterozygous animals, a mouse model of CdLS. Here we show that a significant proportion of the genes whose transcription is altered in brains from the CdLS mouse model contain cohesin in their promoter. We have also found high amounts of cohesin SA1 in the promoter of Zfp608, a zing finger protein whose human ortholog has been described as a biomarker in the clinical identification and classification of CdLS cases. In addition, we show that genes whose expression is altered in these Nipbl or SA1 mutants are located in close proximity in the genome, which suggests a role for SA1 in the architectural organization of such specific loci. This mechanism ensures the coordinated expression of spatially related genes and constitutes an efficient way to co-regulate genes involved in the same biological process. Our results point to a relevant role of cohesin-SA1 in the regulation of genes involved in CdLS pathogenesis.

My participation in this work comprises the design, performance, analysis and interpretation of all experiments, and comments and ideas for the writing of the manuscript.

Cell Cycle 11:12, 2233-2238; June 15, 2012; © 2012 Landes Bioscience

The specific contributions of cohesin-SA1 to cohesion and gene expression

Implications for cancer and development

Ana Cuadrado,[1,*] Silvia Remeseiro,[1] Gonzalo Gómez-López,[2] David G. Pisano[2] and Ana Losada[1]

[1]Chromosome Dynamics Group; Molecular Oncology Programme; Spanish National Cancer Research Centre (CNIO); Madrid, Spain;

[2]Bioinformatics Unit; Structural Biology and Biocomputing Programme; Spanish National Cancer Research Centre (CNIO); Madrid, Spain

Besides its well-established role in sister chromatid cohesion, cohesin has recently emerged as major player in the organization of interphase chromatin. Such important function is related to its ability to entrap two DNA segments also in cis, thereby facilitating long-range DNA looping which is crucial for transcriptional regulation, organization of replication factories and V(D)J recombination. Vertebrate somatic cells have two different versions of cohesin, containing Smc1, Smc3, Rad21/Scc1 and either SA1 or SA2, but their functional specificity has been largely ignored. We recently generated a knockout mouse model for the gene encoding SA1, and found that this protein is essential to complete embryonic development. Cohesin-SA1 mediates cohesion at telomeres, which is required for their replication. Telomere defects in SA1-deficient cells provoke chromosome segregation errors resulting in aneuploidy despite robust centromere cohesion. This aneuploidy could explain why heterozygous animals have an earlier onset of tumorigenesis. In addition, the genome-wide distribution of cohesin changes dramatically in the absence of SA1, and the complex shows reduced accumulation at promoters and CTCF sites. As a consequence, gene expression is altered, leading to downregulation of biological processes related to a developmental disorder linked to cohesin function, the Cornelia de Lange Syndrome (CdLS). These results point out a prominent role of cohesin-SA1 in transcriptional regulation, with clear implications in the etiology of CdLS.

Key words: cohesin, telomere replication, aneuploidy, cancer, chromatin structure, transcription, embryonic development, ChIP-sequencing, CdLS

Submitted: 04/04/12

Accepted: 04/09/12

http://dx.doi.org/10.4161/cc.20318

*Correspondence to: Ana Cuadrado;
Email: acuadrado@cnio.es

Two Different Cohesin Complexes in Somatic Vertebrate Cells

Sister chromatid cohesion is critical to ensure accurate chromosome segregation during cell division and faithful DNA repair by homologous recombination. Such a biological problem is successfully addressed by cohesin, a ring-shaped complex that entraps the DNA fiber and keeps the sister chromatids connected until they part to opposite spindle poles at the metaphase-to-anaphase transition.[1] In addition to this "canonical" role, recent studies have placed cohesin as an important player in other essential biological processes, such as DNA replication,[2] V(D)J recombination[3] and transcriptional regulation (reviewed in ref. 4). In all these instances, cohesin could act by entrapping two DNA segments, either in trans to mediate cohesion, i.e., the sister chromatids, or in cis to stabilize chromatin loops (**Fig. 1A**).

Somatic vertebrate cells have two different versions of cohesin that contain three common subunits, Smc1, Smc3 and Rad21/Scc1, and either stromal antigen 1 (SA1) or SA2, the orthologs of yeast Scc3[5] (**Fig. 1B**). Based on the high degree of homology between the two SA proteins and their similar behavior in terms of association with chromatin during the cell cycle, it has been largely assumed that SA1 and SA2 have redundant functions. Knockdown experiments performed in HeLa cells first suggested that this was

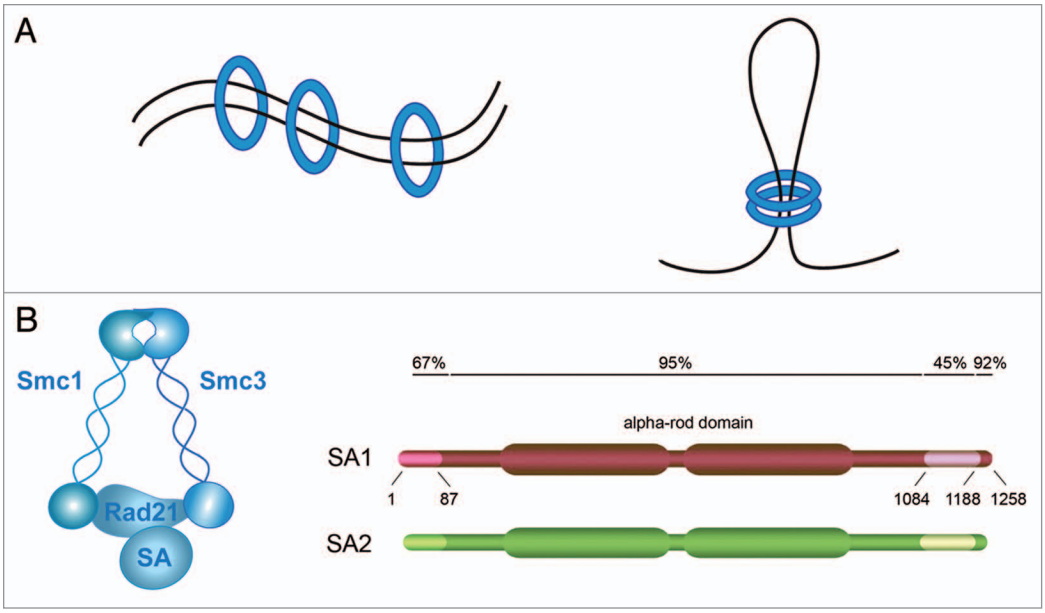


Figure 1. Cohesin composition and mechanism of action. (A) Cohesin (blue circle) can entrap DNA segments in trans to mediate sister chromatid cohesion (left) and in cis to form chromatin loops (right). (B) Subunit composition of cohesin and schematic representation of SA1 and SA2 showing their sequence homology.

not the case by showing that cohesin-SA1 was required for telomere cohesion, while centromere cohesion was performed by cohesin-SA2.[6] The relative contribution of cohesin-SA1 and cohesin-SA2 to other cohesin functions has not been addressed. Importantly, a recent report described the identification of inactivating mutations in the gene encoding SA2, located in the X chromosome, in a significant number of tumors.[7] Loss of SA2 provoked severe cohesion defects leading to aneuploidy in human cancer cell lines. These defects could be rescued after targeted correction of SA2, while gene expression profiles of the corrected cell lines was not significantly different from their SA2-null counterparts. These results suggest that the most important function of SA2 is related to cohesion. In contrast, the characterization of SA1-knockout mice, summarized here, reveals an important contribution of SA1 not only to cohesion, but also to transcriptional regulation.[8,9] This latter function of cohesin has become of special relevance for human disease with the discovery of heterozygous mutations in Nipbl, a factor essential for loading cohesin on chromatin, in half of the patients of Cornelia de Lange syndrome (CdLS).[10,11] This developmental disorder affects 1:10,000–30,000 newborns and is characterized by mental retardation, reduced body size, dysmorphic face, upper limb defects and several additional organ abnormalities.[12] Blood cells from CdLS patients lack obvious cohesion defects, but they display altered expression patterns, featured by minor changes in hundreds of genes.[13] The relevance of these changes in the etiology of the disease remains unclear, and the mechanism(s) by which cohesin causes them are also unknown.

In an effort to determine the specific functions of cohesin-SA1 and cohesin-SA2, we recently generated a knockout mouse model for SA1. Complete loss of SA1 results in embryonic lethality that starts around E11.5, although some late embryos can be obtained.[8] SA1 and SA2 are expressed in most adult mouse tissues, although to different extent. By means of a quantitative analysis, we have estimated that protein levels of SA2 are 3-fold higher than SA1 in MEFs.[8] In wild-type embryos, both SA1 and SA2 show similar distribution by immunohistochemical staining, which does not change for SA2 in the absence of SA1.[9] In some tissues examined, like embryonic brain, the expression of SA2 is also not affected by the lack of SA1. More importantly, the total amount of chromatin-bound cohesin is virtually the same in wild-type and SA1-null cells.[8] Thus, cohesin-SA1 must perform unique roles during mouse development that cannot be fulfilled by cohesin-SA2. One is related to cohesion, whereas the other affects transcription.

Cohesin-SA1, Chromosome Segregation and Cancer

SA1-null MEFs have severe proliferation defects and are aneuploid. We found that cohesin-SA1 and cohesin-SA2 mediate telomere and centromere cohesion, respectively, as proposed in human cells, whereas both contribute to cohesion along the arms (**Fig. 2**). Centromere cohesion is most critical for chromosome segregation, since it must counteract the pulling forces of microtubules during prometaphase to allow proper chromosome alignment.[14] What could then be the role of telomere cohesion and the reason for the aneuploidy in SA1-null cells? Telomeres pose a challenge to the replication machinery, and forks often stall in these regions.[15] Homologous recombination-mediated restart of stalled forks (or repair of DNA breaks generated by fork collapse) probably requires SA1-mediated cohesion so that in cells lacking SA1, telomere replication is inefficient. Incompletely replicated telomeres could be the cause of the lagging chromosomes

and anaphase bridges observed in SA1-null anaphase cells, which often end in either cell death or cytokinesis failure, leading to the formation of tetraploid cells, a known pathway to aneuploidy.[16] Thus, the lack of cohesin-SA1 results in faulty chromosome segregation and drives aneuploidy despite robust centromere cohesion. SA1 heterozygous MEFs show also telomere replication and chromosome segregation defects, although to a lesser extent. Aneuploidy could, therefore, explain our observation that SA1 heterozygous mice show an earlier onset of spontaneous tumorigenesis when compared with their wild-type littermates and, for example, develop pancreatic cancers that are extremely infrequent in mice.[17] Pancreas could be more sensitive to telomere dysfunction and the genomic instability derived from it, as suggested by additional evidence.[18,19] However, we cannot discard a contribution of dysregulated transcription to the tumorigenic effect of SA1 haploinsufficiency.

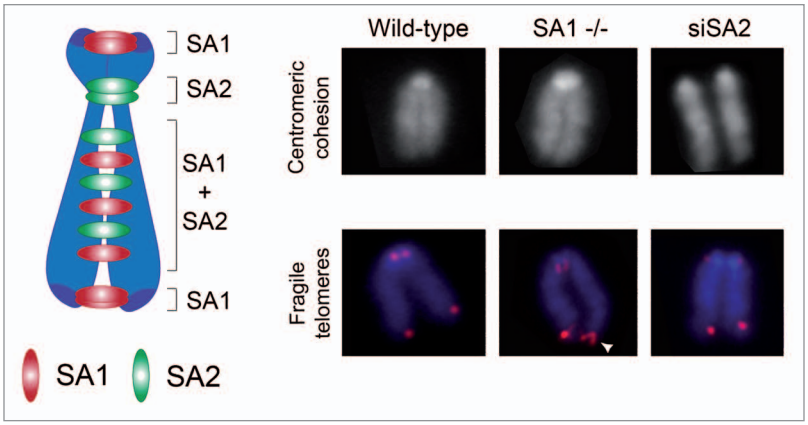


Figure 2. Division of labor between the two cohesin complexes that mediate cohesion along the chromosome. Left: Schematic representation of the distribution of cohesin complexes containing either SA1 and SA2 along the mouse chromosome. Right: Representative images of metaphase chromosomes from MEFs wild-type, SA1-null and wild-type treated with SA2 siRNA stained with DAPI (top) or both with DAPI (blue) and a telomeric repeat FISH probe (red). Centromeric cohesion defects are apparent only in the absence of SA2. Aberrant telomere structures (also known as fragile telomeres) reflect defective replication and are visualized as irregularly shaped telomere signals (white arrowhead).

Cohesin-SA1, Gene Regulation and CdLS

ChIP sequencing analysis revealed remarkable differences in the genomic distributions of cohesin-SA1 and cohesin-SA2.[9] Cohesin-SA1 has greater propensity than cohesin-SA2 to localize at gene-associated regions and, in particular, to gene promoters. More striking, in the absence of SA1, cohesin is driven away from promoters and relocates to intergenic regions. The peak intensity of these new sites is clearly lower, which likely reflects a more random genomic distribution of cohesin in the different cells of the population used for the analysis.[20] In addition, these new positions show reduced colocalization with the chromatin insulator CTCF. In wild-type cells, 71% of the almost 24,000 sites detected with an antibody against SMC1 are also CTCF sites, confirming previous results in mouse and human cells.[21,22] In SA1-null cells, the same antibody detects 47,000 sites, of which 46% are shared with CTCF. If we focus on the cohesin binding sites in SA1-deficient cells distinct from those bound by cohesin SA1 in wild-type cells, the overlap with CTCF drops to 19%. Thus, cohesin-SA2, the only cohesin present in SA1-null cells, lacks the ability to accumulate at promoters and CTCF sites.

The changes in cohesin distribution in SA1-null cells impact transcription, both positively and negatively. Gene ontology (GO)-defined processes related to abnormalities found in CdLS patients are downregulated in SA1-null MEFs. This is consistent with the observation of CdLS features in late SA1-null embryos, such as growth delay, skeletal and bone abnormalities and reduced body size.[9] A highly significant overlap was also obtained when we compared the biological processes altered in SA1-null MEFs and MEFs derived from the Nipbl heterozygous animals, a mouse model of CdLS.[23] In order to explore the relationship between transcriptional changes and cohesin distribution, we selected 130 genes whose expression changes in Nipbl mutant brains, with a fold change (FC) <0.75 or >1.25. Among these genes, 19.2% ($n = 25$) have SA1 binding sites at their promoters (1 kb upstream TSS) according to the ChIP data in MEFs, a percentage much higher than the one observed genome-wide (6.7%, **Fig. 3A**). Similar results were obtained when we compared the percentage of Nipbl dysregulated genes containing SMC1 at their promoters (27.7%) with the genome-wide frequency (10.3%). These data suggest that a significant proportion of the genes regulated in brains from the CdLS mouse model contain cohesin in their promoter [red (SA1) and green (SMC1) lines in the inner part of the circos diagram in **Fig. 3B**]. For some of these genes, like those encoding Protocadherins, we confirmed that cohesin-SA1 is present at their promoters also in the brain, and that reduced transcription in SA1-null embryonic brains correlates with reduced presence of cohesin.[9] Intriguingly, we have found high amounts of cohesin-SA1 in the promoter of Zfp608 (**Fig. 3C and Table 1**), a zinc finger protein whose human ortholog ZNF608 has been recently proposed as a biomarker in the clinical identification and classification of CdLS cases.[13] Little is known about Zfp608, but a report suggests that it could act as a transcription factor regulating the expression of recombination-activating genes (RAG)1 and RAG2 expression during T-cell development.[24] Another transcription factor with a lot of cohesin-SA1 in its promoter is Myc (**Table 2**). Dowregulation of Myc in the absence of cohesin-SA1 is responsible for alterations in the transcription of its targets and likely contributes to the reduced proliferation of SA1-null cells in culture and in embryos.[8,9]

Importantly, presence at the promoter is not the only way in which cohesin regulates gene expression. A survey of the chromosomal location of genes whose expression is altered in either SA1-null

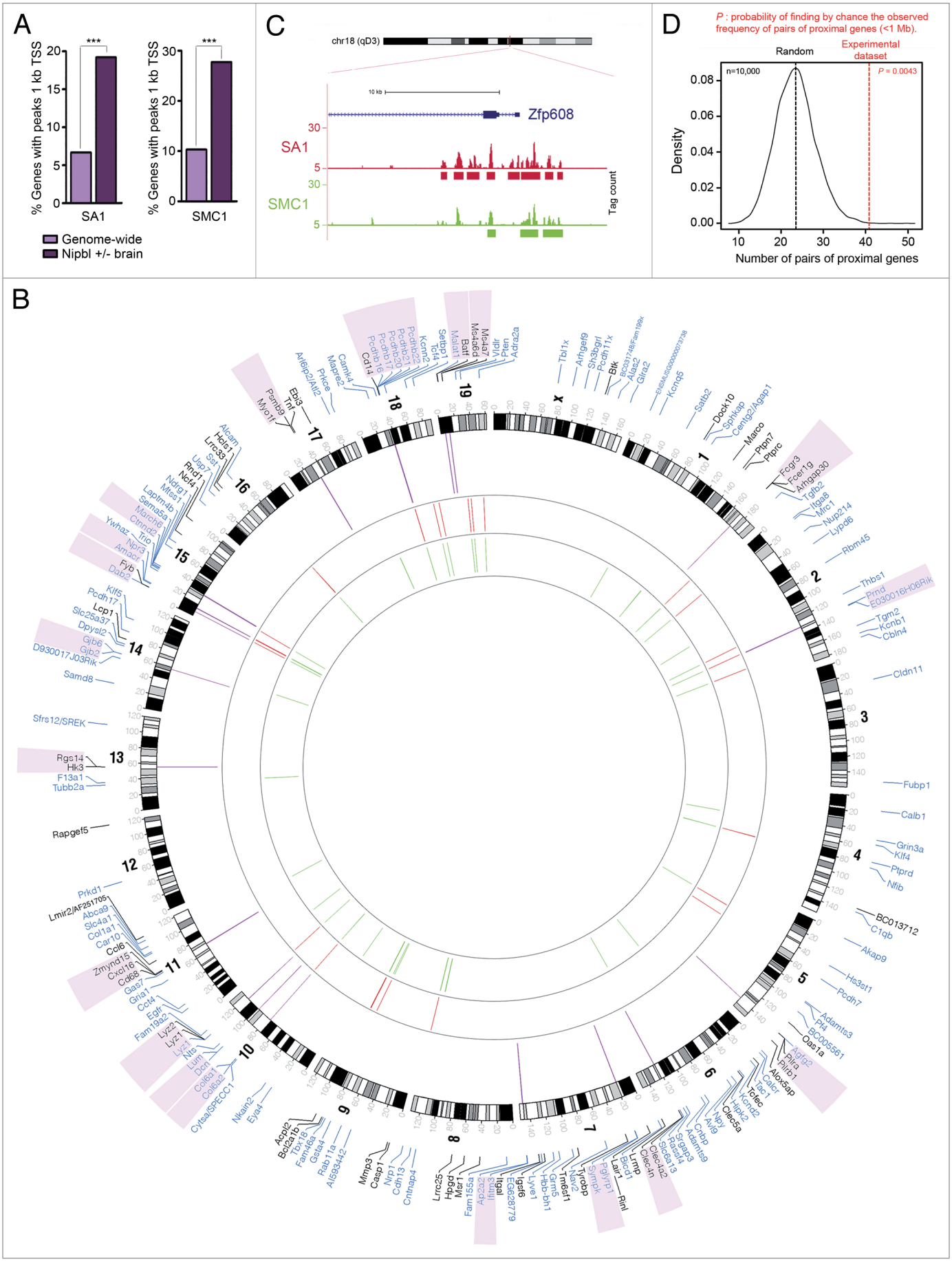

A
% Genes with peaks 1 kb TSS
***
SA1
SMC1
Genome-wide
Nipbl +/- brain
C
chr18 (qD3)
10 kb
Zfp608
SA1
SMC1
Tag count
D
P : probability of finding by chance the observed frequency of pairs of proximal genes (<1 Mb).
Random
Experimental dataset
n=10,000
P = 0.0043
Density
Number of pairs of proximal genes
B

Figure 3 (See opposite page). Transcriptional dysregulation in Nipbl$^{+/-}$ and SA1$^{-/-}$ embryos and CdLS. (A) Percentage of genes with peaks of SA1 or SMC1 located 1 kb upstream Transcription Start Site (TSS) for genes genome-wide or genes dysregulated in Nipbl$^{+/-}$ brains. GraphPad Prism 5 software was used to calculate two-tailed χ^2 test (with Yates' correction, ***p < 0.0001). (B) Circos diagram representing chromosomal location of Differentially Expressed Genes (FDR < 0.15) in SA1-null MEFs (black) and genes dysregulated in brains from Nipbl$^{+/-}$ embryos (blue). Regions shorter than 1 Mb containing two or more regulated genes are indicated by purple lines and the genes are shadowed in the same color. Binding sites at gene promoters (up to 1 kb upstream the TSS) are depicted for SA1 (red) and SMC1 (green). (C) SA1 and SMC1 peaks at Zfp608 gene, a proposed biomarker of CdLS. This is one of the regions in the mouse genome with highest cohesin peak density. (D) Frequency distribution for pairs of proximal genes (< 1 Mb) in 10,000 random data sets of sample size equivalent to our experimental data set (as defined in B, n = 185). Red line highlights the frequency of pairs of proximal genes observed in the experimental data set.

MEFs (differentially expressed genes, FDR < 0.15, n = 55, shown in black in **Fig. 3B**) or Nipbl mutant brains (shown in blue in **Fig. 3B**) reveals that up to one fourth are located nearby (within 1 Mb region) another regulated gene (**Fig. 3B**, purple shadowing and purple lines inside the circos diagram). While in random data sets, the mean of the frequency distribution is 23 pairs of proximal genes, the observed frequency in our experimental data set is 41. The probability of finding by chance such frequency of nearby genes is extremely low (p = 0.0043) (**Fig. 3D**). This suggests a role for SA1 in the architectural organization of such specific loci, which likely accounts for the coordinate expression of clustered genes. This mechanism of control ensures the coordinated expression of spatially related genes and represents an efficient way to co-regulate genes involved in the same biological process.

Thus, our results indicate that cohesin-SA1 has an important role in the regulation of genes involved in CdLS pathogenesis. We speculate that a partial deficiency in Nipbl may restrict cohesin loading. While this restriction would have no major consequences for cohesion, it would significantly affect transcription of genes critical during development. Nevertheless, an involvement of cohesin functions other than gene regulation in CdLS should not be discarded yet. A better characterization of DNA replication, DNA repair and chromosome segregation must be performed in CdLS cells, both from mouse models and human patients.

Outlook

Despite the fact that cohesin was identified 15 years ago, we are still trying to understand how it works. Many aspects of its molecular mechanism of action, its cell cycle regulation or the specific functions of variant cohesin subunits and its associated factors remain a very active field of research. This knowledge will be essential to comprehend the role of somatic and germ line mutations of cohesin in human disease. In the last years, the development of chromosome conformation capture (3C)-based techniques in combination with high-throughput DNA sequencing have much contributed to analyze the organization of chromatin in its natural three-dimensional state. Such techniques are optimal tools to define the cis-regulatory networks responsible for specific transcriptional programs. It is important to highlight that transcription in mammalian cells takes place in large discrete foci, the so-called "transcription factories." In those foci, related genes far away in the genome are organized for efficient and cooperative transcription.[25] We envision that the analysis of the genome-wide distribution of cohesin in different tissues and at different stages of development, together with the characterization of the spatial organization of chromatin derived from such distribution will be key to understand the basis for cell type-specific transcriptional programs. We also expect this information to provide mechanistic explanations for the transcriptional profiles found in diseases in which cohesin function is perturbed, like cancer and CdLS.

Table 1. List of gene promoters with highest density of SA1 binding sites (≥4 SA1 peaks in the 5 kb region upstream the TSS)

Gene symbol	Chromosome	Start	End	SA1 binding sites
Zfp608	chr18	55149834	55154834	4
Zfp949	chr9	88437857	88442857	4
Pcdh7	chr5	58104259	58109259	4
Bdnf	chr2	109528719	109533719	4

Table 2. Top five widest SA1 binding regions

Gene symbol	Chromosome	Start	End	Length (bp)
Myc	chr15	61814652	61820007	5356
Trib2	chr12	15820962	15824328	3367
Hjurp	chr1	90172374	90174983	2610
Sox4	chr13	29043476	29046028	2553
Tsc22d1	chr14	76902752	76905136	2385

All the regions correspond to gene promoters (5 kb upstream the TSS).

Acknowledgments

Research in our lab is supported by the Spanish Ministry of Science and Innovation through grants SAF-2010-21517 and Consolider Ingenio 2010 CSD2007-015 to A.L. and a Ramón y Cajal contract for A.C. S.R. is the recipient of a predoctoral fellowship from "La Caixa."

References

1. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. Annu Rev Genet 2009; 43:525-58; PMID:19886810; http://dx.doi.org/10.1146/annurev-genet-102108-134233.
2. Guillou E, Ibarra A, Coulon V, Casado-Vela J, Rico D, Casal I, et al. Cohesin organizes chromatin loops at DNA replication factories. Genes Dev 2010; 24:2812-22; PMID:21159821; http://dx.doi.org/10.1101/gad.608210.
3. Seitan VC, Krangel MS, Merkenschlager M. Cohesin, CTCF and lymphocyte antigen receptor locus rearrangement. Trends Immunol 2012; 33:153-9; PMID:22440186; http://dx.doi.org/10.1016/j.it.2012.02.004.

4. Dorsett D. Cohesin: genomic insights into controlling gene transcription and development. Curr Opin Genet Dev 2011; 21:199-206; PMID:21324671; http://dx.doi.org/10.1016/j.gde.2011.01.018.
5. Losada A, Yokochi T, Kobayashi R, Hirano T. Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. J Cell Biol 2000; 150:405-16; PMID:10931856; http://dx.doi.org/10.1083/jcb.150.3.405.
6. Canudas S, Smith S. Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells. J Cell Biol 2009; 187:165-73; PMID:19822671; http://dx.doi.org/10.1083/jcb.200903096.
7. Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, Harris BT, et al. Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science 2011; 333:1039-43; PMID:21852505; http://dx.doi.org/10.1126/science.1203619.
8. Remeseiro S, Cuadrado A, Carretero M, Martínez P, Drosopoulos WC, Cañamero M, et al. Cohesin-SA1 deficiency drives aneuploidy and tumorigenesis in mice due to impaired replication of telomeres. EMBO J 2012;31:2076-89; PMID:22415365; http://dx.doi.org/10.1038/emboj.2012.11.
9. Remeseiro S, Cuadrado A, Gómez-López G, Pisano DG, Losada A. A unique role of cohesin-SA1 in gene regulation and development. EMBO J 2012; 31:2090-102; PMID:22415368; http://dx.doi.org/10.1038/emboj.2012.60.
10. Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, et al. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of *Drosophila melanogaster* Nipped-B. Nat Genet 2004; 36:631-5; PMID:15146186; http://dx.doi.org/10.1038/ng1364.
11. Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet 2004; 36:636-41
12. Liu J, Krantz ID. Cornelia de Lange syndrome, cohesin and beyond. Clin Genet 2009; 76:303-14; PMID:19793304; http://dx.doi.org/10.1111/j.1399-0004.2009.01271.x.
13. Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, et al. Transcriptional dysregulation in NIPBL and cohesin mutant human cells. PLoS Biol 2009; 7:1000119; PMID:19468298; http://dx.doi.org/10.1371/journal.pbio.1000119.
14. Vagnarelli P, Morrison C, Dodson H, Sonoda E, Takeda S, Earnshaw WC. Analysis of Scc1-deficient cells defines a key metaphase role of vertebrate cohesin in linking sister kinetochores. EMBO Rep 2004; 5:167-71; PMID:14749720; http://dx.doi.org/10.1038/sj.embor.7400077.
15. Sfeir A, Kosiyatrakul ST, Hockemeyer D, MacRae SL, Karlseder J, Schildkraut CL, et al. Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. Cell 2009; 138:90-103; PMID:19596237; http://dx.doi.org/10.1016/j.cell.2009.06.021.
16. Ganem NJ, Storchova Z, Pellman D. Tetraploidy, aneuploidy and cancer. Curr Opin Genet Dev 2007; 17:157-62; PMID:17324569; http://dx.doi.org/10.1016/j.gde.2007.02.011.
17. Gordon DJ, Resio B, Pellman D. Causes and consequences of aneuploidy in cancer. Nat Rev Genet 2012; 13:189-203; PMID:22269907.
18. Campbell PJ, Yachida S, Mudie LJ, Stephens PJ, Pleasance ED, Stebbings LA, et al. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 2010; 467:1109-13; PMID:20981101; http://dx.doi.org/10.1038/nature09460.
19. van Heek NT, Meeker AK, Kern SE, Yeo CJ, Lillemoe KD, Cameron JL, et al. Telomere shortening is nearly universal in pancreatic intraepithelial neoplasia. Am J Pathol 2002; 161:1541-7; PMID:12414502; http://dx.doi.org/10.1016/S0002-9440(10)64432-X.
20. Carretero M, Remeseiro S, Losada A. Cohesin ties up the genome. Curr Opin Cell Biol 2010; 22:781-7; PMID:20675112; http://dx.doi.org/10.1016/j.ceb.2010.07.004.
21. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell 2008; 132:422-33; PMID:18237772; http://dx.doi.org/10.1016/j.cell.2008.01.011.
22. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature 2008; 451:796-801; PMID:18235444; http://dx.doi.org/10.1038/nature06634.
23. Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, et al. Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome. PLoS Genet 2009; 5:1000650; PMID:19763162; http://dx.doi.org/10.1371/journal.pgen.1000650.
24. Zhang F, Thomas LR, Oltz EM, Aune TM. Control of thymocyte development and recombination-activating gene expression by the zinc finger protein Zfp608. Nat Immunol 2006; 7:1309-16; PMID:17057722; http://dx.doi.org/10.1038/ni1397.
25. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell 2012; 148:84-98; PMID:22265404; http://dx.doi.org/10.1016/j.cell.2011.12.014.

“In questions of science, the authority of a thousand is not worth the humble reasoning of a single individual”.

Galileo Galilei

Discussion

Two different cohesin complexes:

A mouse model to address their functions *in vivo*

The relevance of cohesin for processes other than cohesion, such as transcriptional regulation, replication or recombination, has been increasing rapidly in the last few years. Somatic vertebrate cells have two different versions of cohesin complex that contain either SA1 or SA2. It has been largely considered that both SA proteins have redundant functions, assumption based on their high degree of homology and similar behavior in terms of association to chromatin during the cell cycle. In this work we have generated and characterized a mouse model for cohesin component SA1 and showed that this assumption is incorrect since biallelic inactivation of SA1 results in embryonic lethality. This study sheds light on the functional specificity of each SA protein. In particular, we have addressed the detailed functions of cohesin-SA1 in telomere cohesion and replication together with its role in gene regulation, and explored their implications in human disease, from cancer to Cornelia de Lange Syndrome (CdLS).

Over the following sections, the relevance of the different functions of cohesin complex in human disease will be discussed.

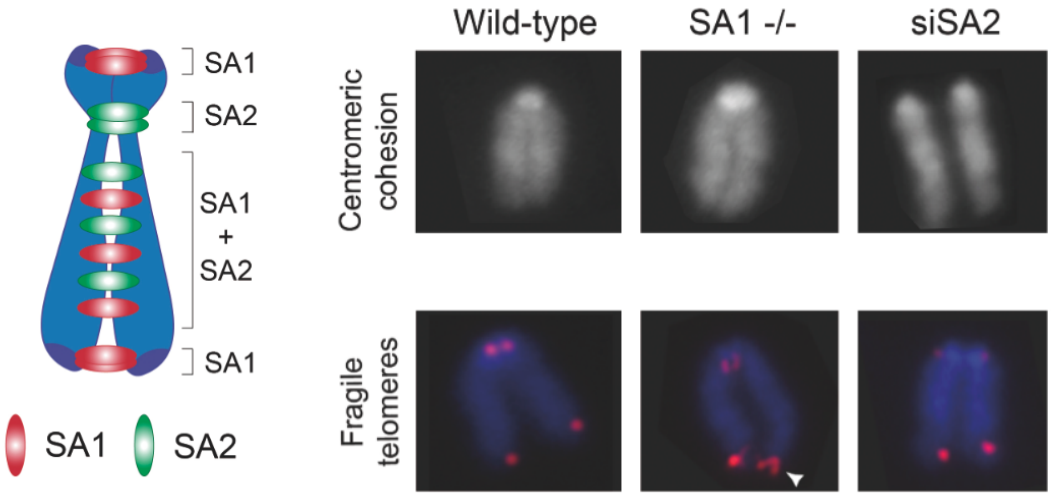
Cohesin-SA1 role in chromosome segregation:

Implications for cancer

Complete ablation of SA1 results in embryonic lethality, which demonstrates that cohesin-SA1 performs essential functions during development that cannot be fulfilled by cohesin-SA2.

Whereas both SA1 and SA2 contribute to cohesion an thereby to prevent fragile site formation along chromosome arms, cohesin-SA2 is critical for centromere cohesion and only cohesin-SA1 can perform this role at telomeres (Figure 8).

Figure 8



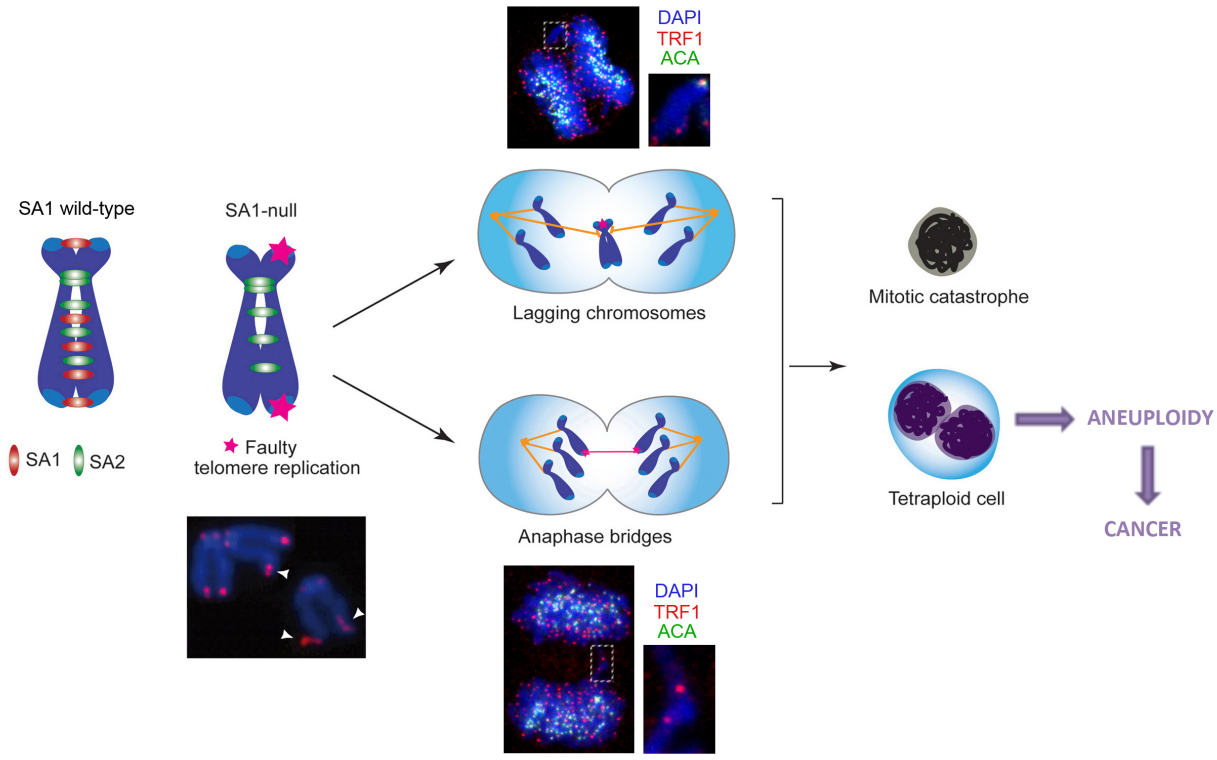
Contribution of cohesin-SA1 and cohesin-SA2 to sister chromatid cohesion at different chromosomal regions: telomeres, centromeres and arms. Left: Scheme of cohesin-SA1 and cohesin-SA2 distribution along the mouse chromosome. Right: Representative images of metaphase chromosomes from wildtype, SA1-null and wildtype cells treated with SA2 siRNA stained with DAPI (top) or both DAPI (blue) and telomere FISH probe (red). Centromere cohesion defects are apparent only in the absence of SA2. Fragile telomeres appear as aberrant telomere structures and reflect defective replication (white arrowhead) upon loss of SA1.

Deficient telomere cohesion in SA1-deficient cells results in defective telomere replication, giving rise to aberrant telomere structures known as fragile telomeres (Sfeir et al. 2009). The mechanistic contribution of cohesin-SA1 to efficient telomere replication involves the cohesive function of cohesin-SA1 since it requires Sororin. Telomeres pose a challenge to the replication machinery (Badie et al. 2010; Ye et al. 2010). Forks often stall in these regions (Gilson and Géli 2007) and, moreover, stalled forks occurring at chromosome ends cannot be rescued by forks progressing in the opposite direction. Fragile telomeres have been also described in mouse cells deficient for components of the shelterin complex that binds the telomeric repeats (Palm and de Lange 2008; Martínez et al. 2009). Such is the case of TRF1 protein that is also required for efficient telomere replication (Sfeir et al. 2009). However, telomere fragility in SA1-null cells does not result from impaired recruitment of shelterin to telomeres, as shown by chromatin fractionation and ChIP-dot blot assays. On other hand, in budding yeast cohesin is critical for the recovery of stalled forks and its loading at the replication sites depends on Rad50, a protein relevant for bridging sister chromatids (Tittel-Elmer et al. 2012). Thus, we propose that cohesin-SA1 ensures efficient telomere replication by stabilizing arrested forks and thereby facilitating their restart and/or by promoting HR-mediated repair of breaks generated upon fork collapse.

Despite their robust centromere cohesion, SA1-null cells present chromosome segregation defects, mainly lagging chromosomes and anaphase bridges, which most likely arise from defective telomere replication. Aberrant telomere structures resulting from incomplete replication in the short arm, next to the centromere, might prevent the separation of sister centromeres, given rise to the observed lagging chromosomes composed of two sister chromatids (Figure 9, top image). When the unresolved replication intermediate happens in the long arm, it might generate instead a chromatin bridge

containing telomeric DNA (Figure 9, bottom image). The presence of these structures prevents proper cytokinesis and results in either mitotic catastrophe or tetraploidization, the latter being a reported initiator of aneuploidy (Ganem et al. 2007; Maser and DePinho 2002; Davoli et al. 2010) (Figure 9).

Figure 9



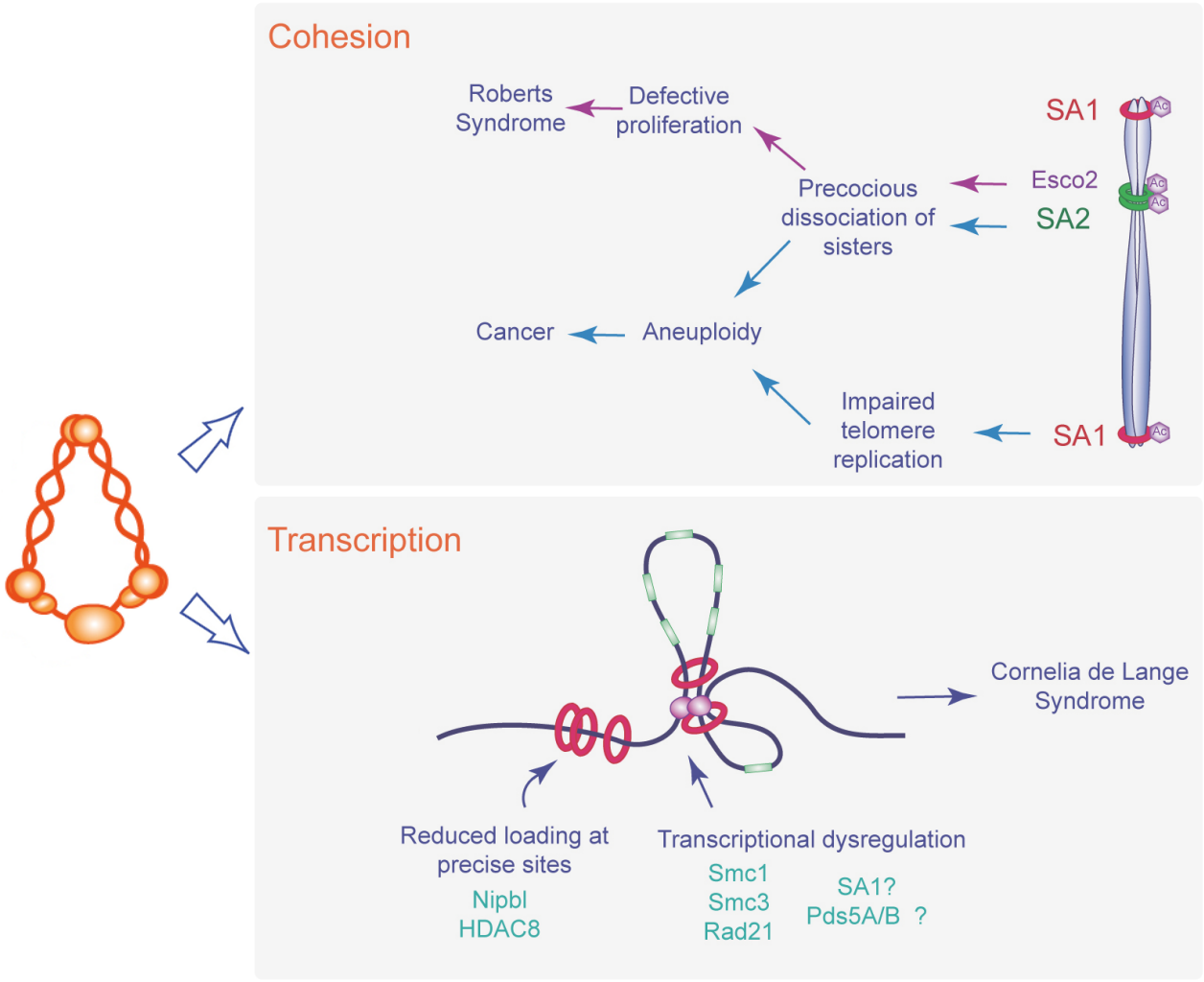
The role of cohesin-SA1 in telomere cohesion and replication is essential for accurate chromosome segregation and to prevent tetraploidization, a driver of aneuploidy and tumourigenesis. Notice the presence of telomeric (TRF1) and not centromeric (ACA) proteins at the anaphase bridges and lagging chromosomes composed of two sister chromatids in SA1-deficient cells.

On other hand, micronuclei haven been largely considered to arise from lagging chromosomes. A mechanism by which errors in mitotic chromosome segregation generates micronuclei and subsequent chromosome pulverization has been proposed and might be

one explanation for chromothripsis (Crasta et al. 2012). This phenomenon, defined by small-scale DNA copy number changes and intrachromosomal rearrangements, has been recently described in human cancer (Stephens et al. 2011; Liu et al. 2011). Two non-exclusive models have been proposed for chromothripsis: 1) chromosome fragmentation followed by stitching together of fragments, and 2) aberrant DNA replication resulting in fork stalling and template switching. Our data might be also consistent with the fact that those complex genomic rearrangements are connected to chromosome catastrophes involving replication mechanisms, although this remains unexplored.

The aneuploidy observed in SA1-deficient cells and tissues likely contributes to tumourigenesis. In fact, SA1 heterozygous mice develop spontaneous tumors earlier than their wildtype littermates and with a broader tumor spectrum. It is remarkable the presence of pancreatic cancers that are extremely infrequent in mice (Gordon et al. 2012). We are currently exploring what is the functional contribution of cohesin-SA1 to pancreatic tumourigenesis. Pancreas might be more sensitive to telomere dysfunction and the genomic instability derived from it than some other organs (Campbell et al. 2010). In order to address this issue, we are employing telomapping approaches, based on the measurement of telomere length (Flores et al. 2008), in tissue samples from pancreas tumors derived from SA1-deficient mice and also in healthy young animals. Our preliminary data suggest that the pancreas from SA1-heterozygous mice tend to have shorter telomeres than their wildtype littermates, which might point to a higher sensitivity to suffer pancreas cancer. However, we cannot rule out a possible contribution of cohesin-SA1 function in regulation of gene expression. In fact, young SA1-heterozygous mice present altered transcription of genes involved in pancreatitis prior to the appearance of tumors, which might predispose them to undergo pancreas tumourigenesis.

Figure 10



Relevance of the different cohesin functions in human disease. (Left) *Cohesion and human disease.* Inactivating mutations in SA2, responsible for centromere cohesion, cause aneuploidy in human cancer due to precocious dissociation of sisters. SA1-deficiency also generates aneuploidy leading to tumourigenesis in mice, but through a different mechanism involving impaired telomere cohesion and replication. Mutations in Esco2, linked to Roberts Syndrome, cause segregation and proliferative defects due to precocious dissociation of sisters. (Right) *Transcription and human disease.* Cornelia de Lange Syndrome is featured by altered transcriptional profiles in the absence of overt cohesion defects. This cohesinopathy has been linked to mutations in the cohesin loader Nipbl and, in few cases, to mutations in cohesin subunits. Impaired cohesin-SA1 role in gene regulation has been proposed to underlie the molecular etiology of CdLS. Pds5A/B-mutant mice present phenotypes reminiscent of CdLS, but their possible implication in transcription is unknown. The contribution of cohesin function in gene expression to cancer remains to be elucidated.

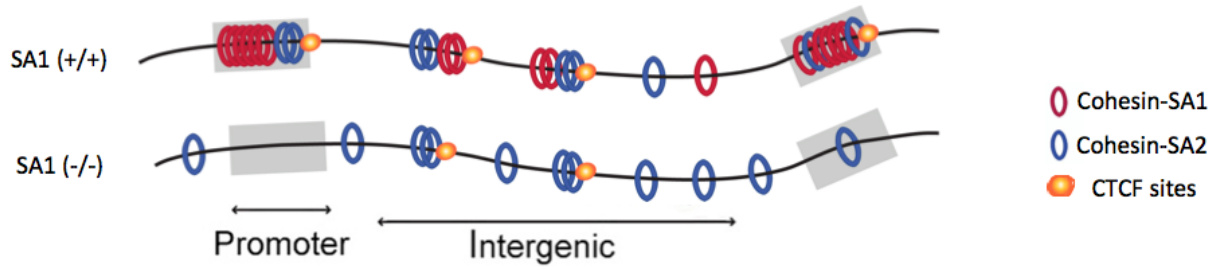
To date, little was known of the involvement of cohesin mutations in human cancer (Figure 10). A recent report described the identification of inactivating mutations in the gene encoding SA2, located in the X chromosome, in a significant number of tumors (Solomon et al. 2011). In those cases, loss of SA2 led to aneuploidy due to precocious separation of sister chromatids in metaphase. Centromere cohesion is most critical for chromosome segregation since it must counteract the pulling forces of microtubules during prometaphase to allow proper chromosome alignment (Vagnarelli et al. 2004; McGuinness et al. 2005). On other hand, cohesin mutations in Smc1, Smc3, Rad21 and SA2 subunits have been found in acute myeloid leukemia, a type of cancer not characterized for being very aneuploidy (Welch et al. 2012). It is unclear how cohesin might contribute to transcription in this case. At least SA2 is not likely to be a major regulator of global gene expression in human cancer, since comparison of transcriptional profiles of cell lines lacking or expressing SA2 does not show significant differences (Solomon et al. 2011). Although we cannot discard a contribution of dysregulated transcription to cancer upon SA1 haploinsufficiency, it is clear that SA1 deficiency generates chromosome segregation defects and aneuploidy, but through a completely different mechanism involving impaired telomere replication in the absence of telomere cohesion, which is mediated exclusively by cohesin-SA1.

Cohesin-SA1 role in gene regulation: Implications in CdLS

Mouse embryos lacking SA1 show developmental delay and die before birth. Their viability strongly decreases by E12.5, but some exceptionally survive to E18.5 and present some features characteristic of CdLS. Genome-wide analysis of cohesin distribution by ChIP-sequencing approaches reveals remarkable differences in the localization of cohesin-SA1

and cohesin-SA2. Whereas cohesin-SA1 has a greater propensity for localizing at gene promoters and gene-associated regions, cohesin-SA2 prefers intergenic regions. More importantly, in the absence of SA1, cohesin-SA2 redistributes to intergenic regions and is driven away from promoters and sites bound by the chromatin insulator CTCF (Figure 11).

Figure 11

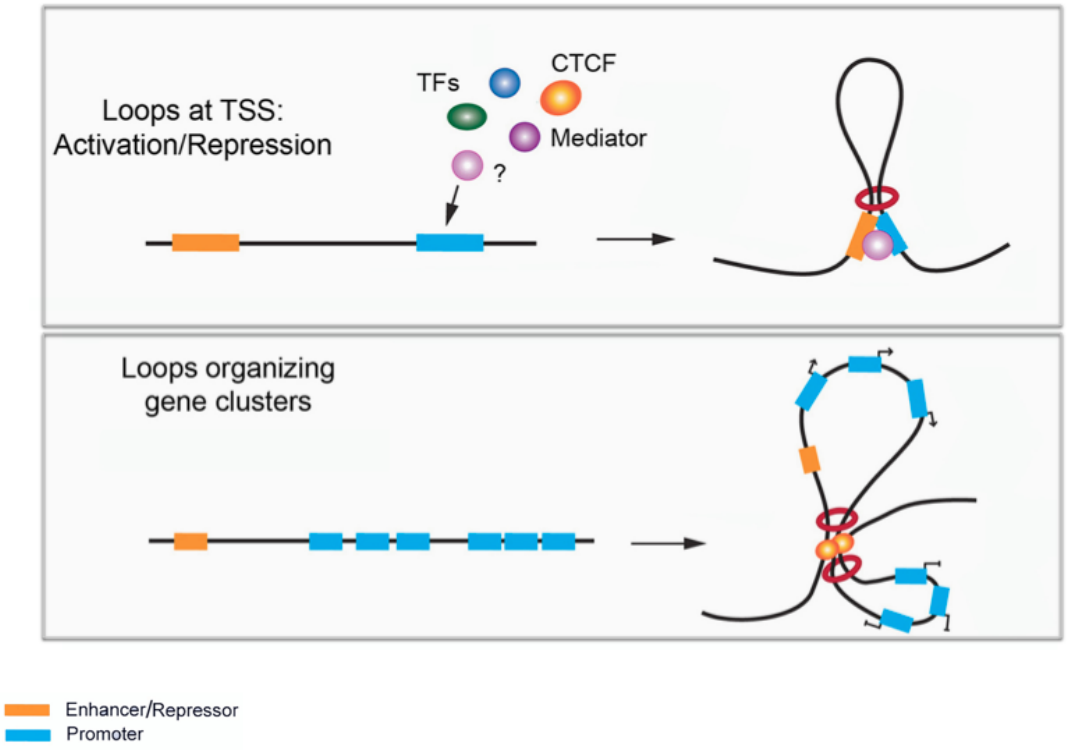


Different dynamics and localization of cohesin-SA1 and cohesion-SA2. SA1 is responsible for cohesin accumulation at promoters and CTCF sites, which clearly points to its role in gene expression regulation. In the absence of SA1, cohesin-SA2 in incapable of replacing SA1 at those sites and, thereby, cohesin is driven away form promoters and CTCF sites. As a consequence, SA1-deficient cells present an altered transcriptional profile.

The number of binding sites for Smc1 in wildtype and SA1-null MEFs doubles in the absence of SA1. The new sites to which cohesin is relocated are characterized by reduced colocalization with CTCF and low cohesin occupancy, which likely reflects a more random genomic distribution of cohesin. Therefore, cohesin-SA2 is incapable of accumulating at promoters and CTCF sites when SA1 is not present. As a consequence of these changes, transcription is altered in the absence of SA1.

The mechanisms underlying the role of cohesin-SA1 in regulation of gene expression are likely related to the ability of the complex to bring together DNA sequences in *cis* facilitating DNA looping (Hadjur et al. 2009; Mishiro et al. 2009; Nativio et al. 2009; Kagey et al. 2010; Hou et al. 2010; Chien et al. 2011). Chromatin looping causally underlies gene regulation (Deng et al. 2012), therefore loops may determine gene expression by promoting or preventing the interaction between promoters and enhancers in different modes. Some configurations would require cohesin presence at promoters, while some others would determine the organization of a gene cluster (Figure 12).

Figure 12



Proposed mechanisms for cohesin-SA1 in gene expression regulation. Cohesin-SA1 presence at promoters (top) or in the vicinity of genes organized in clusters (bottom) is required for the formation of loops that arrange the chromatin for gene transcription. CTCF as well as different transcription factors (TFs) are likely involved in cohesin-SA1 recruitment to specific genomic positions.

The changes in cohesin distribution in SA1-null cells impact transcription, indicating that cohesin-SA2 cannot fulfill the function of cohesin-SA1 in terms of gene expression regulation. We hypothesize that SA1 and SA2 must confer distinct properties to the cohesin complex, most likely making cohesin-SA1 more prone to occupy fixed genomic positions and cohesin-SA2 more prone to slide along DNA. Importantly, these different features would have a differential impact on cohesin functions: cohesion, the most important function of cohesin-SA2, does not require cohesin localization at precise sites, whereas regulation of transcription does. It remains to be elucidated the molecular mechanism underlying these potential differences in mobility between cohesin-SA1 and cohesin-SA2. It could be related to either a differential interaction between SA proteins and the cohesin regulatory factors Pds5 and Wapl or to specific interactions with proteins located at defined sites. We and others have observed that cohesin-SA1 and cohesin-SA2 can both interact with CTCF, as judged by immunoprecipitation assays (Xiao et al. 2011). Future analysis using quantitative proteomics may help us elucidate specific protein-protein interactions for SA1 and SA2.

The increasing interest in cohesin regulating gene expression arises from the identification of mutations in the cohesin loader Nipbl and cohesin subunits as a major cause of CdLS (Figure 10). Cells derived from those patients do not show cohesion defects (Liu and Krantz 2009; Dorsett 2011), but present altered transcriptional profiles. These features together with some other pathologies observed in CdLS patients, such as craniofacial anomalies, heart defects or delayed ossification, are recapitulated in a mouse model heterozygous for a gene-trap mutation in Nipbl gene (Kawauchi et al. 2009). SA1 heterozygous mice do not show apparent CdLS phenotypes, but SA1-null embryos surviving up to late embryogenesis stages do present features that resemble CdLS pathologies. More importantly, gene ontology (GO)-defined processes related to abnormalities found in CdLS

patients are downregulated in SA1-null cells, and this transcriptional changes significantly overlap with the altered profiles of Nipbl heterozygous MEFs. In addition, a significant proportion of the genes dysregulated in brains from the CdLS mouse model contain cohesin-SA1 in their promoter (e.g. myc, protocadherins) which points to a clear relationship between transcriptional changes and altered cohesin-SA1 function in this human syndrome. Protocadherins (Pcdhs) are cell adhesion molecules involved in tissue morphogenesis and establishment / maintenance of selective neuronal connections during development and adulthood (Monahan et al. 2012). Most genes encoding Pcdhs are located in three consecutive clusters (a, b, g), being a and g regulated in a complex way involving alternative promoter choice followed by alternative *cis* splicing. Pcdhs genes have been found downregulated in the brains of Nipbl-heterozygous mice, leading the authors to propose that reduced levels of protocadherins in the brain might contribute to the mental retardation observed in CdLS patients (Kawauchi et al. 2009). We have shown that regulation of Pcdh gene expression in the brain relies on the presence of cohesin-SA1 at their promoters.

However, the presence at promoters is not the only way in which cohesin regulates gene expression. We found that cohesin-SA1 determines cohesin distribution within gene clusters and that genes whose expression is altered in SA1-null cells and Nipbl mutant brains are located nearby in the genome. This points to a role for SA1 in the architectural organization of such loci, which provides a mechanism for controlling and ensuring coordinated expression of spatially related genes and co-regulation of genes involved in the same biological processes.

We thus propose that cohesin-SA1 plays an important role in the regulation of genes involved in the etiology of CdLS. Partial deficiency in Nipbl might affect differently the binding of cohesin-SA1 and cohesin-SA2; either because SA1 is less abundant (Losada et al. 2000; Holzmann et al. 2011) or because it is more often associated to genes and it needs to

be constantly reloaded by Nipbl to allow passage of the transcriptional machinery. We hypothesize that reduction of Nipbl levels may decrease the loading of both cohesin complexes, but while this barely affects cohesion, it does affect transcriptional regulation of genes that are critical during development.

All this, together with additional studies, points to the relevance of the dynamic behavior of the complex in order to achieve a threshold of cohesin at sites in which the presence of the complex is critical for regulated gene expression. This has become more evident after identification of HDAC8 mutations in a subset of CdLS patients (Deardorff et al. 2012a). This cohesin dynamic behavior must be modulated by its loader, by its interaction with Pds5/Wapl, its acetylation state (and thereby Esco1/2 and HDAC8) and by its association with factors like CTCF (through SA/Scc3) (Figure 10), although the underlying mechanistic features remain to be elucidated.

Perspectives

Currently, cohesin research is a very active field and the interest in different aspects of its biology has been increasing more and more over the last few years. In fact, despite of being discovered fifteen years ago for its role in sister chromatid cohesion, many features of its mechanism of action, its regulation and the particular functions of variant cohesin subunits and its associated factors remain to be clarified.

The work presented in this thesis indicates that cohesin-SA1 plays a relevant role in telomere cohesion and replication with clear implications in cancer, but it is also essential for the accurate regulation of genes involved in CdLS pathogenesis. However, many underlying aspects have not been dissected yet. Therefore, we should not discard a contribution of transcriptional dysregulation to the tumourigenic effect of SA1

haploinsufficiency and the involvement of some cohesin functions other than gene regulation in CdLS.

On one hand, it is not clear if Nipbl loads cohesin-SA1 and cohesin-SA2 in a slightly different manner or if partial deficiency in Nipbl affects differently the cohesive and the transcriptional functions of the complex (Figure 10). In fact, we are currently pursuing this line of research and our preliminary data suggest that reduced Nipbl levels has no major impact on the bulk loading of the two cohesin complexes, on DNA repair and on chromosome segregation. However, limiting amounts of Nipbl affect the loading of cohesin at certain gene promoters, which in turn impacts its transcription. Thus, reduced Nipbl levels restrict cohesin loading just locally, and while cohesion is unaffected, transcriptional regulation of critical genes during development is altered.

On other hand, recent reports evidence the relevance of chromatin structure in determining gene expression (Deng et al. 2012). The development of Chromosome Conformation Capture (3C) technology and the subsequent variants (Dekker et al. 2002; Splinter et al. 2012), together with Next Generation Sequencing, have contributed to the analysis of chromatin organization from a genome-wide three-dimensional perspective. The spatial organization of transcription domains, in which genes are arranged for efficient and coordinated transcription, can be studied by these kind of genome-wide approaches (Dixon et al. 2012; Shen et al. 2012).

We are currently employing similar strategies in specific tissues and at different developmental stages in order to dissect the contribution of SA1 to the establishment of their transcriptomes. We have observed that cohesin shows specific distributions in different tissues and at different stages. In particular, cohesin is tissue-specifically enriched at the promoters of the genes being expressed in each tissue. We are also correlating active and repressive histone marks with cohesin distribution in each tissue or developmental

stage, and employing Hidden Markov Models to describe cohesin-defined chromatin states. We anticipate that integrating data from the genome-wide distribution of cohesin in different tissues and developmental stages, together with their transcriptional profiles and information of the spatial organization of chromatin, will be key to understand how cohesin defines chromatin states and how tissue-specific transcriptional programs are established. More importantly, this information will also help to better understand the mechanisms underlying human diseases, in which cohesin functions are altered, such as cancer and CdLS.

“I think and think for months and years.
Ninety-nine times, the conclusion is false.
The hundredth time I am right”.
Albert Einstein

Conclusions

1. Ablation of SA1 results in late embryonic lethality in mouse; thus, some functions of cohesin-SA1 that cannot be assumed by cohesin-SA2 are essential to fulfill embryonic development.

2. Cohesion at telomeres is mediated specifically by cohesin-SA1 and contributes to their efficient replication. Cohesin-SA2 is responsible for centromere cohesion, whereas both perform this function at chromosome arms.

3. Defective telomere replication causes chromosome segregation defects in SA1-deficient cells, despite their robust centromere cohesion, and leads to aneuploidy. This mechanism of generating aneuploidy in the absence of cohesin-SA1 is different from the classical one involving precocious sister chromatid separation.

4. SA1 heterozygous animals have increased risk of cancer, but are protected against acute tumourigenesis induced by chemical carcinogens.

5. SA1-null embryos that reach late embryonic stages present a clear growth delay and some features reminiscent of CdLS.

6. Genome-wide cohesin distribution, and in particular its accumulation at promoters and CTCF sites, depends largely on the presence of SA1 which points to a unique role of cohesin-SA1 in transcriptional regulation.

7. SA1-null cells present transcriptional changes related to biological functions altered in CdLS. Impaired cohesin-SA1 function in gene expression may underlie the molecular etiology of this human syndrome.

8. Cohesin-SA1 regulates gene expression by different mechanisms, one involves its presence at gene promoters (e.g. myc, protocadherins) and the other entails the spatial organization of gene clusters. Lack of SA1 affects the binding of cohesin at promoters and along gene clusters and, thereby, leads to altered transcription since cohesin-SA2 cannot assume those functions.

Conclusiones

1. La eliminación de SA1 provoca letalidad embrionaria. La cohesina-SA2 no puede asumir ciertas funciones de la cohesina-SA1 que son esenciales para el desarrollo embrionario.

2. La cohesión de los telómeros está mediada específicamente por cohesina-SA1 y es necesaria para que la replicación ocurra de forma eficiente. La cohesina-SA2 es responsable de la cohesión centromérica, mientras que ambos complejos cooperan para llevar a cabo esta función a lo largo de los brazos cromosómicos.

3. Una replicación ineficiente de los telómeros en las células deficientes en SA1 causa defectos de segregación cromosómica, a pesar de que presentan una cohesión centromérica robusta, y provoca aneuploidías. Este mecanismo de generación de aneuploidía en ausencia de cohesina-SA1 es diferente del mecanismo clásico caracterizado por una separación precoz de las cromátidas hermanas.

4. Los ratones heterocigotos para el gen SA1 presentan mayor incidencia de tumores espontáneos, pero están protegidos ante la inducción aguda de tumores mediante tratamiento con carcinógenos químicos.

5. Los embriones deficientes en SA1 que alcanzan estadios tardíos del desarrollo embrionario presentan un claro retraso en el crecimiento y algunas características propias del síndrome de Cornelia de Lange.

6. La distribución genómica de cohesina, y en particular su acumulación en promotores y sitios CTCF, depende en gran medida de la presencia de SA1, lo cual apunta claramente a una función única de SA1 en regulación transcripcional.

7. Las células deficientes en SA1 presentan cambios transcripcionales relacionados con procesos biológicos que están alterados en los pacientes CdLS. El papel de la cohesina-SA1 en regulación de la expresión génica guarda relación directa con la etiología de CdLS.

8. La cohesina-SA1 regula la expresión génica mediante distintos mecanismos, uno implica su presencia en promotores génicos (ej. myc, protocaderinas) y otro consiste en la organización espacial de *clústeres* génicos. La ausencia de SA1 afecta a la unión de cohesina en promotores y a lo largo de *clústeres* y, por tanto, conlleva una alteración de la transcripción dado que cohesina-SA2 no puede asumir esas funciones.

References

Almedawar S, Colomina N, Bermúdez-López M, Pociño-Merino I, Torres-Rosell J. 2012. A SUMO-Dependent Step during Establishment of Sister Chromatid Cohesion. *Curr Biol*. (In press).

Badie S, Escandell JM, Bouwman P, Carlos AR, Thanasoula M, Gallardo MM, Suram A, Jaco I, Benitez J, Herbig U, et al. 2010. BRCA2 acts as a RAD51 loader to facilitate telomere replication and capping. *Nature Structural & Molecular Biology* **17**: 1461–1469.

Barber TD, McManus K, Yuen KWY, Reis M, Parmigiani G, Shen D, Barrett I, Nouhi Y, Spencer F, Markowitz S, et al. 2008. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci USA* **105**: 3443–3448.

Beckouët F, Hu B, Roig MB, Sutani T, Komata M, Uluocak P, Katis VL, Shirahige K, Nasmyth K. 2010. An Smc3 acetylation cycle is essential for establishment of sister chromatid cohesion. *Mol Cell* **39**: 689–699.

Borges V, Lehane C, Lopez-Serra L, Flynn H, Skehel M, Rolef Ben-Shahar T, Uhlmann F. 2010. Hos1 deacetylates Smc3 to close the cohesin acetylation cycle. *Mol Cell* **39**: 677–688.

Brough R, Bajrami I, Vatcheva R, Natrajan R, Reis-Filho JS, Lord CJ, Ashworth A. 2012. APRIN is a cell cycle specific BRCA2-interacting protein required for genome integrity and a predictor of outcome after chemotherapy in breast cancer. *EMBO J* **31**: 1160–1176.

Campbell PJ, Yachida S, Mudie LJ, Stephens PJ, Pleasance ED, Stebbings LA, Morsberger LA, Latimer C, McLaren S, Lin M-L, et al. 2010. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**: 1109–1113.

Canudas S, Smith S. 2009. Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells. *The Journal of Cell Biology* **187**: 165–173.

Caron P, Aymard F, Iacovoni JS, Briois S, Canitrot Y, Bugler B, Massip L, Losada A, Legube G. 2012. Cohesin protects genes against γH2AX Induced by DNA double-strand breaks. *PLoS Genet* **8**: e1002460.

Castronovo P, Gervasini C, Cereda A, Masciadri M, Milani D, Russo S, Selicorni A, Larizza L. 2009. Premature chromatid separation is not a useful diagnostic marker for Cornelia de Lange syndrome. *Chromosome Res* **17**: 763–771.

Chan K-L, Roig MB, Hu B, Beckouët F, Metson J, Nasmyth K. 2012. Cohesin's DNA Exit Gate Is Distinct from Its Entrance Gate and Is Regulated by Acetylation. *Cell*. (In press).

Chien R, Zeng W, Kawauchi S, Bender MA, Santos R, Gregson HC, Schmiesing JA, Newkirk DA, Kong X, Ball AR, et al. 2011. Cohesin mediates chromatin interactions that regulate mammalian β-globin expression. *J Biol Chem* **286**: 17870–17878.

Covo S, Westmoreland JW, Gordenin DA, Resnick MA. 2010. Cohesin Is limiting for the suppression of DNA damage-induced recombination between homologous chromosomes. *PLoS Genet* **6**: e1001006.

Crasta K, Ganem NJ, Dagher R, Lantermann AB, Ivanova EV, Pan Y, Nezi L, Protopopov A, Chowdhury D, Pellman D. 2012. DNA breaks and chromosome pulverization from errors in mitosis. *Nature* **482**: 53–58.

Davoli T, Denchi EL, de Lange T. 2010. Persistent Telomere Damage Induces Bypass of Mitosis and Tetraploidy. *Cell* **141**: 81–93.

Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, Saitoh K, Komata M, Katou Y, Clark D, et al. 2012a. HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature*. (In press).

Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodríguez C, Arnedo M, Loeys B, Kline AD, et al. 2007. Mutations in Cohesin Complex Members SMC3 and SMC1A Cause a Mild Variant of Cornelia de Lange Syndrome with Predominant Mental Retardation. *The American Journal of Human Genetics* **80**: 485–494.

Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, Mönnich M, Yan Y, Xu W, Gil-Rodríguez MC, et al. 2012b. RAD21 Mutations Cause a Human Cohesinopathy. *Am J Hum Genet* **90**: 1014–1027.

Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, Vettermann C, Lin YC, Ju Z, Schulz D, et al. 2011. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci USA* **108**: 9566–9571.

Degner SC, Wong TP, Jankevicius G, Feeney AJ. 2009. Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J Immunol* **182**: 44–48.

Dekker J, Rippe K, Dekker M, Kleckner N. 2002. Capturing chromosome conformation. *Science* **295**: 1306–1311.

Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, Dean A, Blobel GA. 2012. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping

factor. *Cell* **149**: 1233–1244.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376–380.

Dorsett D. 2011. Cohesin: genomic insights into controlling gene transcription and development. *Current Opinion in Genetics & Development* **21**: 199–206.

Dreier MR, Bekier ME, Taylor WR. 2011. Regulation of sororin by Cdk1-mediated phosphorylation. *Journal of Cell Science* **124**: 2976–2987.

Fay A, Misulovin Z, Li J, Schaaf CA, Gause M, Gilmour DS, Dorsett D. 2011. Cohesin selectively binds and regulates genes with paused RNA polymerase. *Curr Biol* **21**: 1624–1634.

Flores I, Canela A, Vera E, Tejera A, Cotsarelis G, Blasco MA. 2008. The longest telomeres: a general signature of adult stem cell compartments. *Genes & Development* **22**: 654–667.

Gandhi R, Gillespie PJ, Hirano T. 2006. Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Curr Biol* **16**: 2406–2417.

Ganem NJ, Storchova Z, Pellman D. 2007. Tetraploidy, aneuploidy and cancer. *Current Opinion in Genetics & Development* **17**: 157–162.

Gilson E, Géli V. 2007. How telomeres are replicated. *Nat Rev Mol Cell Biol* **8**: 825–838.

Glynn EF, Megee PC, Yu H-G, Mistrot C, Unal E, Koshland DE, DeRisi JL, Gerton JL. 2004. Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. *Plos Biol* **2**: E259.

Gordon DJ, Resio B, Pellman D. 2012. Causes and consequences of aneuploidy in cancer. *Nat Rev Genet* **13**: 189–203.

Gruber S, Arumugam P, Katou Y, Kuglitsch D, Helmhart W, Shirahige K, Nasmyth K. 2006. Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. *Cell* **127**: 523–537.

Guacci V, Koshland D, Strunnikov A. 1997. A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. cerevisiae. *Cell* **91**: 47–57.

Guillou E, Ibarra A, Coulon V, Casado-Vela J, Rico D, Casal I, Schwob E, Losada A, Méndez J.

2010. Cohesin organizes chromatin loops at DNA replication factories. *Genes & Development* **24**: 2812–2822.

Gullerova M, Proudfoot NJ. 2008. Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. *Cell* **132**: 983–995.

Gutiérrez-Caballero C, Cebollero LR, Pendás AM. 2012. Shugoshins: from protectors of cohesion to versatile adaptors at the centromere. *Trends Genet* **28**: 351–360.

Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M. 2009. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature*. **460**: 410-413.

Haering CH, Farcas A-M, Arumugam P, Metson J, Nasmyth K. 2008. The cohesin ring concatenates sister DNA molecules. *Nature* **454**: 297–301.

Heidinger-Pauli JM, Mert O, Davenport C, Guacci V, Koshland D. 2010. Systematic reduction of cohesin differentially affects chromosome segregation, condensation, and DNA repair. *Curr Biol* **20**: 957–963.

Herrán Y, Gutiérrez-Caballero C, Sánchez-Martín M, Hernández T, Viera A, Barbero JL, de Álava E, de Rooij DG, Suja JÁ, Llano E, et al. 2011. The cohesin subunit RAD21L functions in meiotic synapsis and exhibits sexual dimorphism in fertility. *EMBO J* **30**: 3091–3105.

Holzmann J, Fuchs J, Pichler P, Peters J-M, Mechtler K. 2011. Lesson from the stoichiometry determination of the cohesin complex: a short protease mediated elution increases the recovery from cross-linked antibody-conjugated beads. *J Proteome Res* **10**: 780–789.

Hou C, Dale R, Dean A. 2010. Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc Natl Acad Sci USA* **107**: 3651–3656.

Hu B, Itoh T, Mishra A, Katoh Y, Chan K-L, Upcher W, Godlee C, Roig MB, Shirahige K, Nasmyth K. 2011. ATP hydrolysis is required for relocating cohesin from sites occupied by its Scc2/4 loading complex. *Curr Biol* **21**: 12–24.

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**: 430–435.

Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, et al. 2009. Multiple Organ System Defects and Transcriptional

Dysregulation in the Nipbl+/− Mouse, a Model of Cornelia de Lange Syndrome ed. W.A. Bickmore. *PLoS Genet* **5**: e1000650.

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, et al. 2004. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. *Nat Genet* **36**: 631–635.

Kueng S, Hegemann B, Peters BH, Lipp JJ, Schleiffer A, Mechtler K, Peters J-M. 2006. Wapl controls the dynamic association of cohesin with chromatin. *Cell* **127**: 955–967.

Kurze A, Michie KA, Dixon SE, Mishra A, Itoh T, Khalid S, Strmecki L, Shirahige K, Haering CH, Löwe J, et al. 2011. A positively charged channel within the Smc1/Smc3 hinge required for sister chromatid cohesion. *EMBO J* **30**: 364–378.

Lafont AL, Song J, Rankin S. 2010. Sororin cooperates with the acetyltransferase Eco2 to ensure DNA replication-dependent sister chromatid cohesion. *Proc Natl Acad Sci USA* **107**: 20364–20369.

Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, Itoh T, Watanabe Y, Shirahige K, Uhlmann F. 2004. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**: 573–578.

Liu J, Krantz ID. 2009. Cornelia de Lange syndrome, cohesin, and beyond. *Clin Genet* **76**: 303–314.

Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, et al. 2009. Transcriptional Dysregulation in NIPBL and Cohesin Mutant Human Cells ed. N. Hastie. *Plos Biol* **7**: e1000119.

Liu P, Erez A, Nagamani SCS, Dhar SU, Kołodziejska KE, Dharmadhikari AV, Cooper ML, Wiszniewska J, Zhang F, Withers MA, et al. 2011. Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell* **146**: 889–903.

Losada A. 2005. Functional contribution of Pds5 to cohesin-mediated cohesion in human cells and Xenopus egg extracts. *Journal of Cell Science* **118**: 2133–2141.

Losada A, Hirano M, Hirano T. 1998. Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. *Genes & Development* **12**: 1986–1997.

Losada A, Yokochi T, Kobayashi R, Hirano T. 2000. Identification and characterization of

SA/Scc3p subunits in the Xenopus and human cohesin complexes. *The Journal of Cell Biology* **150**: 405–416.

Martínez P, Thanasoula M, Muñoz P, Liao C, Tejera A, McNees C, Flores JM, Fernández-Capetillo O, Tarsounas M, Blasco MA. 2009. Increased telomere fragility and fusions resulting from TRF1 deficiency lead to degenerative pathologies and increased cancer in mice. *Genes & Development* **23**: 2060–2075.

Maser RS, DePinho RA. 2002. Connecting chromosomes, crisis, and cancer. *Science* **297**: 565–569.

McAleenan A, Cordon-Preciado V, Clemente-Blanco A, Liu I-C, Sen N, Leonard J, Jarmuz A, Aragon L. 2012. SUMOylation of the α-Kleisin Subunit of Cohesin Is Required for DNA Damage-Induced Cohesion. *Curr Biol*.

McGuinness BE, Hirota T, Kudo NR, Peters J-M, Nasmyth K. 2005. Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. *Plos Biol* **3**: e86.

Merkenschlager M. 2010. Cohesin: a global player in chromosome biology with local ties to gene regulation. *Current Opinion in Genetics & Development* **20**: 555–561.

Michaelis C, Ciosk R, Nasmyth K. 1997. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91**: 35–45.

Milutinovich M, Unal E, Ward C, Skibbens RV, Koshland D. 2007. A multi-step pathway for the establishment of sister chromatid cohesion. *PLoS Genet* **3**: e12.

Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M. 2009. Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO J* **28**: 1234–1245.

Misulovin Z, Schwartz YB, Li X-Y, Kahn TG, Gause M, MacArthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, et al. 2008. Association of cohesin and Nipped-B with transcriptionally active regions of the Drosophila melanogaster genome. *Chromosoma* **117**: 89–102.

Monahan K, Rudnick ND, Kehayova PD, Pauli F, Newberry KM, Myers RM, Maniatis T. 2012. Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of Protocadherin-α gene expression. *Proc Natl Acad Sci USA* **109**: 9125–9130.

Mönnich M, Kuriger Z, Print CG, Horsfield JA. 2011. A zebrafish model of Roberts syndrome reveals that Esco2 depletion interferes with development by disrupting the cell cycle.

PLoS ONE **6**: e20051.

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L. 2006. X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. *Nat Genet* **38**: 528–530.

Muto A, Calof AL, Lander AD, Schilling TF. 2011. Multifactorial origins of heart and gut defects in nipbl-deficient zebrafish, a model of Cornelia de Lange Syndrome. *Plos Biol* **9**: e1001181.

Nasmyth K, Haering CH. 2009. Cohesin: its roles and mechanisms. *Annu Rev Genet* **43**: 525–558.

Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters J-M, Murrell A. 2009. Cohesin Is Required for Higher-Order Chromatin Conformation at the Imprinted IGF2-H19 Locus ed. W.A. Bickmore. *PLoS Genet* **5**: e1000739.

Nishiyama T, Ladurner R, Schmitz J, Kreidl E, Schleiffer A, Bhaskara V, Bando M, Shirahige K, Hyman AA, Mechtler K, et al. 2010. Sororin mediates sister chromatid cohesion by antagonizing Wapl. *Cell* **143**: 737–749.

Oliveira RA, Hamilton RS, Pauli A, Davis I, Nasmyth K. 2010. Cohesin cleavage and Cdk inhibition trigger formation of daughter nuclei. *Nat Cell Biol* **12**: 185–192.

Palidwor GA, Shcherbinin S, Huska MR, Rasko T, Stelzl U, Arumughan A, Foulle R, Porras P, Sanchez-Pulido L, Wanker EE, et al. 2009. Detection of alpha-rod protein repeats using a neural network and application to huntingtin. *PLoS Comput Biol* **5**: e1000304.

Palm W, de Lange T. 2008. How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**: 301–334.

Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, et al. 2008. Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. *Cell* **132**: 422–433.

Pauli A, van Bemmel JG, Oliveira RA, Itoh T, Shirahige K, van Steensel B, Nasmyth K. 2010. A direct role for cohesin in gene regulation and ecdysone response in Drosophila salivary glands. *Curr Biol* **20**: 1787–1798.

Revenkova E, Eijpe M, Heyting C, Hodges CA, Hunt PA, Liebe B, Scherthan H, Jessberger R. 2004. Cohesin SMC1 beta is required for meiotic chromosome dynamics, sister

chromatid cohesion and DNA recombination. *Nat Cell Biol* **6**: 555–562.

Rolef Ben-Shahar T, Heeger S, Lehane C, East P, Flynn H, Skehel M, Uhlmann F. 2008. Eco1-dependent cohesin acetylation during establishment of sister chromatid cohesion. *Science* **321**: 563–566.

Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. 2008. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci USA* **105**: 8309–8314.

Schaaf CA, Misulovin Z, Sahota G, Siddiqui AM, Schwartz YB, Kahn TG, Pirrotta V, Gause M, Dorsett D. 2009. Regulation of the Drosophila Enhancer of split and invected-engrailed gene complexes by sister chromatid cohesion proteins. *PLoS ONE* **4**: e6202.

Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT. 2010. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Research* **20**: 578–588.

Schmitz J, Watrin E, Lénárt P, Mechtler K, Peters J-M. 2007. Sororin is required for stable binding of cohesin to chromatin and for sister chromatid cohesion in interphase. *Curr Biol* **17**: 630–636.

Schüle B, Oviedo A, Johnston K, Pai S, Francke U. 2005. Inactivating mutations in ESCO2 cause SC phocomelia and Roberts syndrome: no phenotype-genotype correlation. *Am J Hum Genet* **77**: 1117–1128.

Seitan VC, Banks P, Laval S, Majid NA, Dorsett D, Rana A, Smith J, Bateman A, Krpic S, Hostert A, et al. 2006. Metazoan Scc4 homologs link sister chromatid cohesion to cell and axon migration guidance. *Plos Biol* **4**: e242.

Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, et al. 2011. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* **476**: 467–471.

Sfeir A, Kosiyatrakul ST, Hockemeyer D, MacRae SL, Karlseder J, Schildkraut CL, de Lange T. 2009. Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. *Cell* **138**: 90–103.

Shamu CE, Murray AW. 1992. Sister chromatid separation in frog egg extracts requires DNA topoisomerase II activity during anaphase. *The Journal of Cell Biology* **117**: 921–934.

Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV,

et al. 2012. A map of the cis-regulatory sequences in the mouse genome. *Nature*. (In press).

Shintomi K, Hirano T. 2010. Sister chromatid resolution: a cohesin releasing network and beyond. *Chromosoma* **119**: 459–467.

Sjögren C, Nasmyth K. 2001. Sister chromatid cohesion is required for postreplicative double-strand break repair in Saccharomyces cerevisiae. *Current Biology* **11**: 991–995.

Sjögren C, Ström L. 2010. S-phase and DNA damage activated establishment of sister chromatid cohesion--importance for DNA repair. *Experimental Cell Research* **316**: 1445–1453.

Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, Harris BT, Toretsky JA, Rosenberg SA, Shukla N, Ladanyi M, et al. 2011. Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science* **333**: 1039–1043.

Splinter E, de Wit E, van de Werken HJG, Klous P, de Laat W. 2012. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: From fixation to computation. *Methods*. (In press).

Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. 2008. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J* **27**: 654–666.

Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, et al. 2011. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**: 27–40.

Sumara I, Vorlaufer E, Gieffers C, Peters BH, Peters JM. 2000. Characterization of vertebrate cohesin complexes and their regulation in prophase. *The Journal of Cell Biology* **151**: 749–762.

Terret M-E, Sherwood R, Rahman S, Qin J, Jallepalli PV. 2009. Cohesin acetylation speeds the replication fork. *Nature* **462**: 231–234.

Tittel-Elmer M, Lengronne A, Davidson MB, Bacal J, François P, Hohl M, Petrini JHJ, Pasero P, Cobb JA. 2012. Cohesin Association to Replication Sites Depends on Rad50 and Promotes Fork Restart. *Mol Cell*. (In press).

Tonkin ET, Wang T-J, Lisgo S, Bamshad MJ, Strachan T. 2004. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in

Cornelia de Lange syndrome. *Nat Genet* **36**: 636–641.

Unal E, Heidinger-Pauli JM, Kim W, Guacci V, Onn I, Gygi SP, Koshland DE. 2008. A molecular determinant for the establishment of sister chromatid cohesion. *Science* **321**: 566–569.

Vagnarelli P, Morrison C, Dodson H, Sonoda E, Takeda S, Earnshaw WC. 2004. Analysis of Scc1-deficient cells defines a key metaphase role of vertebrate cohesin in linking sister kinetochores. *EMBO Rep* **5**: 167–171.

Vega H, Waisfisz Q, Gordillo M, Sakai N, Yanagihara I, Yamada M, van Gosliga D, Kayserili H, Xu C, Ozono K, et al. 2005. Roberts syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion. *Nat Genet* **37**: 468–470.

Wang LH-C, Schwarzbraun T, Speicher MR, Nigg EA. 2008. Persistence of DNA threads in human anaphase cells suggests late completion of sister chromatid decatenation. *Chromosoma* **117**: 123–135.

Watrin E, Peters J-M. 2009. The cohesin complex is required for the DNA damage-induced G2/M checkpoint in mammalian cells. *EMBO J* **28**: 2625–2635.

Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters J-M. 2006. Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Curr Biol* **16**: 863–874.

Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC, Wartman LD, Lamprecht TL, Liu F, Xia J, et al. 2012. The Origin and Evolution of Mutations in Acute Myeloid Leukemia. *Cell* **150**: 264–278.

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, et al. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**: 796–801.

Whelan G, Kreidl E, Wutz G, Egner A, Peters J-M, Eichele G. 2012. Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin. *EMBO J* **31**: 71–82.

Xiao T, Wallace J, Felsenfeld G. 2011. Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Molecular and Cellular Biology* **31**: 2174–2183.

Xu H, Balakrishnan K, Malaterre J, Beasley M, Yan Y, Essers J, Appeldoorn E, Tomaszewski JM, Thomaszewski JM, Vazquez M, et al. 2010. Rad21-cohesin haploinsufficiency impedes DNA repair and enhances gastrointestinal radiosensitivity in mice. *PLoS ONE* **5**: e12112.

Xu H, Beasley MD, Warren WD, van der Horst GTJ, McKay MJ. 2005. Absence of mouse REC8 cohesin promotes synapsis of sister chromatids in meiosis. *Dev Cell* **8**: 949–961.

Xu H, Tomaszewski JM, McKay MJ. 2011. Can corruption of chromosome cohesion create a conduit to cancer? *Nat Rev Cancer* **11**: 199–210.

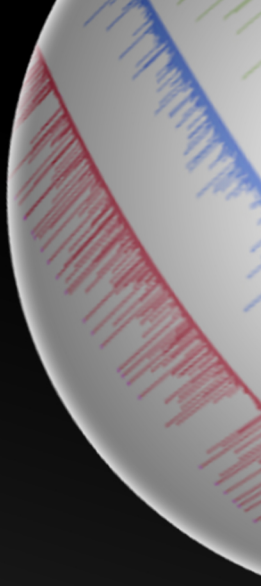
Ye J, Lenain C, Bauwens S, Rizzo A, Saint-Léger A, Poulet A, Benarroch D, Magdinier F, Morere J, Amiard S, et al. 2010. TRF2 and apollo cooperate with topoisomerase 2alpha to protect human telomeres from replicative damage. *Cell* **142**: 230–242.

Zhang B, Chang J, Fu M, Huang J, Kashyap R, Salavaggione E, Jain S, Kulkarni S, Shashikant K, Deardorff MA, et al. 2009. Dosage effects of cohesin regulatory factor PDS5 on mammalian development: implications for cohesinopathies. *PLoS ONE* **4**: e5232.

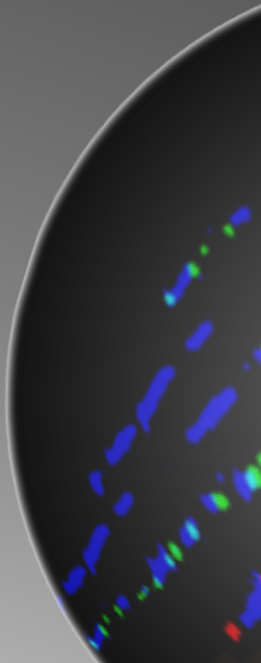
Zhang B, Jain S, Song H, Fu M, Heuckeroth RO, Erlich JM, Jay PY, Milbrandt J. 2007. Mice lacking sister chromatid cohesion protein PDS5B exhibit developmental abnormalities reminiscent of Cornelia de Lange syndrome. *Development* **134**: 3191–3201.

Zhang J, Shi X, Li Y, Kim B-J, Jia J, Huang Z, Yang T, Fu X, Jung SY, Wang Y, et al. 2008a. Acetylation of Smc3 by Eco1 is required for S phase sister chromatid cohesion in both human and yeast. *Mol Cell* **31**: 143–151.

Zhang N, Kuznetsov SG, Sharan SK, Li K, Rao PH, Pati D. 2008b. A handcuff model for the cohesin complex. *The Journal of Cell Biology* **183**: 1019–1031.



"If you're not failing every now and again,
it's a sign you're not doing anything very innovative".
Woody Allen



Doctoral Thesis performed at the Chromosome Dynamics Group
Centro Nacional de Investigaciones Oncológicas (CNIO)
Madrid, 2012